In the past quarter century, research has proliferated on self-regulation and, especially, on self-regulated learning (SRL). Cataloging and synthesizing this literature is now a substantial task, the scope of which is powerfully evidenced by the number and variety of chapters in this handbook. Both basic and applied research need to wrestle with questions about measuring constructs associated with SRL, including components such as metacognition, motivation, and strategic action. Measuring SRL and its components is the topic of this chapter.

Our chapter is divided into four sections. In the first section, we develop key points about measurement in general. In the second section, we describe a model of SRL and its components that researchers and practitioners seek to measure. Next, we survey protocols currently used to measure SRL, including questionnaires, structured interviews, teacher ratings, think aloud methods, error detection tasks, trace methodologies, and observations. In a final section, we constructively critique current protocols for measuring SRL and forecast what measurements of SRL might be like in the future.
1. MEASURING INTERVENES IN AN ENVIRONMENT

Measuring, except for genuinely unobtrusive measurements (Webb, Campbell, Schwartz, & Sechrest, 1966) and some traces (Winne, 1982), intervenes in a student's environment. We design measurement instruments with an intention to cause the student to recall or to generate a particular kind of response. In this view, instruments for measuring SRL and its components—a questionnaire item, a request to stop and talk about current thoughts, a chapter placed before a student with instructions to make any desired marks on pages of the chapter while studying it for a subsequent test—are, in principle, equivalent to an independent variable in an experiment. Interpreting measurement data, therefore, must attend to the same issues as those related to making valid interpretations of causal factors in general.

Cook and Campbell's (1979) typology of these issues has four main categories. First, issues of statistical conclusion validity concern factors that affect whether covariation between the intervention created by measuring and a student's response (a) can be detected, in the sense of having sufficient statistical power to avoid a false negative conclusion by wrongly declaring the absence of covariation between the measurement intervention and response because of a rare or inadequate sample, and (b) whether covariation has been falsely detected, in the sense of having sufficient statistical confidence to avoid claiming a relationship between the measurement intervention and response that is a false positive. Second, issues of internal validity concern the so-called third variable problem or confounding between the measurement intervention, other factors, and a response. When a response is given to a particular measuring instrument, is the response attributable to that intervention or to some other factor in the student's environment? Third, construct validity of putative causes and putative effects is a set of concerns about whether the instruments, as they are operationally defined, represent the components we intend to measure and not other components, and, whether the response reflects what we intend it to reflect. Finally, external validity or generalizability refers to factors that affect the degree to which the particular intervention and the specific observed response are representative of those in other circumstances or environments.

Messick's (1989) typology is another powerful set of lenses through which to view measurements as interventions. He focuses on issues of construct validity under two main headings. One is whether a construct is underrepresented in the sense that its components are not appropriately included in the measuring instrument and the environment in which that instrument is used to intervene. The other is whether the observed data reflect components that are irrelevant to the focal component.
At least two significant implications follow from characterizing measures of SRL as interventions in an environment. First, to adapt a maxim, no instrument is an island. Care should be exercised to map the environment in which measuring instruments intervene and the nature of the disturbance an instrument makes in that environment. We use this "intrusive" language deliberately because it reminds us that disturbances are, as Messick notes, issues to be investigated as experiments. When we measure, we change the environment.

This first inference leads directly to a second. It is that developers of measuring instruments must recognize they are operationally defining a theory of SRL in the instrument they develop. This is the topic of the next section on modeling SRL.

II. MEASUREMENTS OF SELF-REGULATED LEARNING REFLECT A MODEL OF SELF-REGULATED LEARNING

As the history of science shows, theory typically advances in reciprocal and recursive interaction with work to "engineer" measures related to theory. We believe this same interaction characterizes research and measurement of SRL. From the prior section, we also argue that theorizing should attend to factors of the environment into which measuring instruments are inserted because, if environmental factors are disturbed by measurement interventions, those effects may bear on internal validity and generalizability. Obviously, this complicates theorizing and, as well, work on developing measures of SRL.

A. COMPONENTS OF SELF-REGULATED LEARNING

The term "self-regulated" is associated with forms of learning that are metacognitively guided, at least partly intrinsically motivated, and strategic (Winne, 1995, 1997; Zimmerman, 1990). Metacognition is the awareness learners have about their general academic strengths and weaknesses, cognitive resources they can apply to meet the demands of particular tasks, and their knowledge about how to regulate engagement in tasks to optimize learning processes and outcomes. Intrinsic motivation refers to self-regulated learners' belief in incremental learning, a high value placed on personal progress and deep understanding as opposed to besting peers or impressing others, high efficacy for learning, and attributions that link outcomes to factors under their control (e.g., effective use of strategies). "Strategic" describes the way in which these learners approach challenging tasks and problems by choosing from a repertoire of tactics those they believe are best suited to the situation, and applying those tactics appropri-
ately. From this description, it is plain that many facets of SRL are not readily observable. Therefore, one challenge in studying SRL is to find ways to document its components.

SRL has properties of an aptitude and an event (Winne, 1997; Winne & Stockley, 1998). An aptitude describes a relatively enduring attribute of a person that predicts future behavior. For example, a student answering a simple question about whether she or he adapts studying to the circumstances of assignments might say "Yes." On this basis, we might predict the student would study for tomorrow's quiz differently than when preparing an outline for a term project. Moreover, it should not matter whether the question is asked when the work is handed in or a week before, and it should not matter whether the quiz is about history and the project is on biology or vice versa.

An event is like a snapshot that freezes activity in motion, a transient state embedded in a larger, longer series of states unfolding over time. Although our everyday sense of time is one of an unbroken flow, measured time is not continuous. At fundamental physical levels and at levels that characterize SRL, time happens in discrete chunks. As a temporal entity, an event therefore has a beginning and an end; for example, an endpoint in an oscillating crystal or a decision to try a different tactic when studying does not seem to be progressing well enough that is made following deliberation. The event prior to the current one is like a brief-lived aptitude that may predict the next event. Two main features that differentiate aptitude measures of SRL from event-related measures of SRL are discussed throughout the next two sections: aggregation and the kind of information represented in the measurement.

B. SELF-REGULATED LEARNING AS APITUDE

When SRL is measured as an aptitude, a single measurement aggregates over or abstracts some quality of SRL based on multiple SRL events. For instance, in the context of a student studying for a quiz, a researcher may be interested in metacognitive monitoring concerning a rehearsal tactic. The student's description of self-regulation might be recorded in several forms: a rating on a questionnaire item, an interviewer's classification of the student's response to a probe about what activities were used to study for the quiz, or the proportion of particular kinds of notes written in a textbook chapter. In each case, the student or the researcher abstracts over multiple events of monitoring to characterize the set of those events in a single datum generated in response to the measurement intervention. Measurements of SRL as aptitude often are used to predict whether a student will or will not, can or can not act on an SRL-related cognition such as a belief about performing a study tactic. In this sense, a measure-
**TABLE 1 Components of Self-Regulated Learning as Aptitude**

<table>
<thead>
<tr>
<th>Metacognitive Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Knowledge of fine-grained cognitive operations that comprise cognitive tactics</td>
</tr>
<tr>
<td>- Knowledge about strategies that articulate cognitive tactics</td>
</tr>
<tr>
<td>- Procedural knowledge that enacts cognitive tactics</td>
</tr>
<tr>
<td>- Conditional knowledge about occasions to enact cognitive tactics</td>
</tr>
<tr>
<td>- Knowledge of tasks' parameters (e.g., resources, standards for success)</td>
</tr>
<tr>
<td>- Knowledge of self-parameters (e.g., interest, effort)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metacognitive Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Difficulty in addressing the task (ease of learning, EOL)</td>
</tr>
<tr>
<td>- Match of achievement to standards (judgments of learning, JOL)</td>
</tr>
<tr>
<td>- Probability of retrieval from long-term memory (feeling of knowing, FOK)</td>
</tr>
<tr>
<td>- Confidence about the accuracy of monitoring</td>
</tr>
</tbody>
</table>

The measurement of SRL as aptitude is unary—it can stand alone, independent of other measurements.

The most common protocols for measuring SRL as an aptitude include questionnaires and structured interviews, although teacher ratings have been used also (Zimmerman & Martinez-Pons, 1988). When measured as an aptitude, SRL varies within individuals over relatively long time periods, within individuals across different tasks and settings, and across individuals (Pintrich, Wolters, & Baxter, in press; Winne, 1996). In Table 1 (modeled after Pintrich, Wolters, & Baxter, in press) we list facets of SRL as aptitude that have been examined in contemporary research.

**C. SELF-REGULATED LEARNING AS EVENT**

An event spans time, but is marked off along a timeline by a prior event and a subsequent event. We propose that SRL measured as an event has three successively more complex levels: occurrence, contingency, and patterned contingency.

An occurrence of SRL as event is observed when there is a transition from a first state, where an observable feature indicating SRL was not in existence, to a second state, where an indicator of SRL is present. For instance, suppose a student is solving a geometry problem and thinking aloud as the task unfolds. The student might muse aloud, "Wow, this is hard!" To have made that judgment, we presume the student monitored the current state of the task. Every monitoring event must use standards as points against which to monitor, and the standards used here might have been the number of steps taken so far, the time taken to reach this point in the task, or a revised estimate of the probability that the goal or endpoint of the task will be achieved (Winne, 1997). To count the student's report as SRL requires inferences that (a) one or more of those scales for judging task difficulty exceeded a threshold that had not been crossed.
previously, (b) the student thereby classified this task as "difficult," and (c) said so as a consequence of these cognitions. Note that the occurrence of metacognitive monitoring was not directly measured, but the student's report is interpreted as indirect evidence that it occurred.

SRL as a contingency has the form of a binary conditional relationship. It is often modeled in if–then form. Some data describe each of the ifs and other data describe the then. For example, if we observe the student who exclaimed, "Wow, this is hard!," to then take out note paper and make a note, there are two measurements—one of metacognitive monitoring of the difficulty of a task in progress and a second of metacognitive control exercised when the student made the note. This if–then contingency is an elemental cognitive tactic. Over repeated occurrences, the transition from the prior state—monitoring difficulty—to the subsequent state—exercising metacognitive control—can be quantified by a probability that the if transitions to a particular then. A probability of 1.0 indicates that the prior if is always followed by a specific then state. A transition probability less than 1.0 indicates the if state is followed by at least two then states. Whereas probabilities must sum to unity, the relative likelihood of each can be known. If a transition probability is less than 1.0 and only one then state is represented in the observer's measurement system, the system is incomplete, measurements of the if and then states are unreliable, or both.

SRL as a patterned contingency assembles several singular if–then contingencies into a structured ensemble (Corno, 1993; Winne, 1995, 1996, 1997). For example, several basic cognitive tactics might be arrayed as a cognitive strategy involving decision making (monitoring) that selects (controls) which particular tactics are applied in a task as a function of on-line feedback the student generates or receives from external sources (Butler & Winne, 1995; Winne, 1995). Graphic representations of SRL as a patterned contingency might resemble flow charts or path models (Winne, Gupta, & Nesbit, 1994), where the arrows or lines are labeled with probabilities of transitioning from a prior state to a subsequent one.

D. THE SELF-REGULATED LEARNING MODEL OF WINNE AND HADWIN

A full view of SRL as an event spans three and sometimes four phases (Winne & Hadwin, 1997). Within a phase, cognitive operations, labeled as the Operation(s) box in Figure 1, construct a particular kind of product. Four kinds of products, enumerated in the box labeled Products, differentiate the four phases. In this model, information can play one of four roles: a condition, a product, an evaluation, or a standard. The arrows in Figure 1 represent paths along which information flows to update conditions, products, evaluations, and standards. Also depicted in Figure 1 are two events critical to SRL: metacognitive monitoring and metacognitive control.
Phase 1. Defining the Task

In Phase 1, the learner generates perceptions of the task at hand. These perceptions may be simple portraits, but they are more likely elaborate, multifaceted, personalized blends of information attended to in the environment plus memories about similar tasks experienced previously (Butler & Winne, 1995; Winne, 1997). There are two sources of information that contribute to these perceptions in Winne and Hadwin's model. Task conditions provide information about the task in the environment "as it
is.” Cognitive conditions are memorial representations of some features of similar past tasks. These features might include knowledge of the domain(s) of the task (e.g., spelling, searching the Internet), memories about self in relation to the task (e.g., attributions for qualities of prior achievements, interest), and memories about tactics and strategies applied in previous encounters with the same or similar tasks.

Cognitive operations such as searching memory work on conditions to create a definition of the task, thereby creating the first product listed in the Products box in Figure 1. We hypothesize that students generate at least two products in Phase 1. One is a default perception that characterizes the task if it is addressed using a standard or routine approach, whatever the student views that to be. The second perception is a hypothesis about what the task will be like if a nonstandard approach is taken. This hypothesis seems necessary on logical grounds: Even if there is just one tactic the student perceives as standard for the task, the student can exercise agency in Phase 3 to engage that tactic or to dismiss it (see Winne, 1997); the latter being the nonstandard approach. We conjecture that recognizing the existence of choices like this is one hallmark of SRL.

Phase 1 definitions of a task can be metacognitively monitored relative to standards. In Figure 1, five such standards are listed in the Standards box, abstractly labeled A, B, C, D, and E. Suppose the student is reading a chapter in an ecology text that has proved to be very demanding to understand and that standard B refers to amount of domain knowledge. This standard derives from a cognitive condition, a belief, the student holds. The height of the bar for standard B indicates the student believes that a lot of domain knowledge is required when studying this text.

The first sentence in the chapter, “Recall the model of the water cycle presented in the last chapter,” is a task condition. By applying cognitive operations to this information, the student defines this task in terms of the domain knowledge it entails. This is represented by the tall bar associated with B in the Product box. If this product is monitored against its corresponding standard, the cognitive evaluation that results is a judgment about how equipped the student is to approach the task as he or she defined it, represented in the box of Cognitive Evaluations, where factor B is described as on target. If this were the only facet of the task considered, the student probably would exercise cognitive control to continue to Phase 2 because the perception generated about the task seems accurate enough to proceed.

Tasks are likely represented by more than a single facet of information, however. Other facets of this task—A, C, D, and E—constitute the student’s full definition of the task. Monitoring the match of each of these products against their respective standards may lead to other forms of metacognitive control that recycle through Phase 1 to redefine the task. For example, suppose standard D refers to epistemological beliefs about
the worth of spending effort to study ecology from the text. The student’s initial standard for this was zero, but information available in the task’s presentation, perhaps an indication of how significant the water cycle is to understanding conservation of rain forests, leads him or her to define the task as one involving a strong correlation between effort and worth. The student may be drawn to exercise metacognitive control in a form that reevaluates the task’s conditions and to develop a new definition of the task. Similar flows of information occur within and between other phases of Winne and Hadwin’s model.

Phase 2. Setting Goals and Planning How to Reach Them

Cognition in Phase 2 is decision making, supplemented by information retrieved from memory, to frame goals and assemble a plan for approaching them. In Figure 1, goals appear in the Standards box and are modeled as multivariate profiles of information (Butler & Winne, 1995; Winne & Hadwin, 1998). Each standard that makes up a goal profile is a value against which processes and products of engagement can be monitored. By cycling through this stage, the student may affirm precursors to goals that are initially developed in Phase 1 or the student may update a goal for the task that will have been generated during engagement with the task’s central work, reserved for Phase 3. Once goals are active, memory may automatically retrieve tactics or strategies coupled to them (McKoon & Ratcliff, 1992). Plans constituted this way are a common sign of expertise.

As was the case with Phase 1, activities in Phase 2 may involve several cycles “around” the model when cognitive evaluations are not on target and lead the student to reframe cognitive conditions, or seek to change task conditions.

Phase 3. Enacting Tactics

Applying tactics and strategies identified in Phase 2 marks a transition into Phase 3, where tactics copy information into or construct information in working memory. Tactics are bundles of memories comprised of conditional knowledge (ifs that characterize a tactic’s appropriateness) and cognitive operations (thens that change and construct information). Contemporary theory suggests that conditional knowledge consists of two classes of information. Cold propositions describe what a tactic is and does. Hot propositions—efficacy expectations, outcome expectations, incentives associated with completing (or failing to complete) a task, and attributions—are motivational beliefs that give rise or link to affect (Pintrich, Marx, & Boyle, 1993; Winne, 1995, 1997; Winne & Marx, 1989).

The cognitive and behavioral products that tactics create also are modeled as a multivariate profile. When the student monitors products, this generates internal feedback. If the products of Phase 3 are translated into overt behavior, external feedback also may be available from the
environment when, say, a computer command fails to work or a peer remarks on qualities of a student's answer to a question.

**Phase 4. Adapting Metacognition**

Optionally, in Phase 4, the student makes major adaptations to those parts of the model under the student's control. This is accomplished (after Rumelhart & Norman, 1978) in three ways: by accreting (or deleting) conditions under which operations are carried out or by changing operations themselves, by tuning features that account for how tactics articulate as an event, or by significantly restructuring cognitive conditions, tactics, and strategies to create very different approaches to addressing tasks (Winne, 1997).

Although Figure 1 may imply that SRL unfolds as a linear sequence of phases, this is not likely the case. SRL is recursive in two ways. First, as illustrated for Phase 1, information generated in a particular phase may feed back into that same phase if metacognitive monitoring identifies a difference between the current state of cognition and a standard (goal). Second, the product of monitoring in a later phase may feed into a prior phase, as when a judgment about how well a "tried and true" tactic worked in Phase 3 invites the student to recalibrate initial perceptions about the task's difficulty or complexity available from Phase 1. In this sense, SRL is weakly sequenced. Once a first perception of a task is generated in Phase 1, information generated in that and any subsequent phase may jump phases or recurse to the same phase.

**The Centrality of Monitoring and Feedback**

**In Self-Regulated Learning**

Metacognitive monitoring is the gateway to self-regulating one's learning (Butler & Winne, 1995; Winne, 1996, 1997) because without the cognitive evaluations it creates, there is no standard against which to enact regulation. Like other cognitive operations, monitoring produces information, a list of matches and mismatches between (a) the standards for a task and (b) a representation in working memory of the product(s) of (a phase of) a task. As illustrated earlier, when a student develops a definition of a task in Phase 1, that perception may be monitored relative to memories about similar prior tasks. In Phase 3, monitoring operations contrast the products that have been created so far in the task against standards (goals) framed in Phase 2. The products of monitoring, labeled cognitive evaluations in Figure 1, can update to any or all of (a) task conditions, for example, if the student is prompted to ask a question of the teacher that leads to a change in resources provided for the task, (b) cognitive conditions, for instance, when a student revises motivational elements that bear on the task, (c) standards (goals) the student assigns to the task, and (d) products generated in a first cycle of cognition or on the basis of recursive
cognition, for instance, changing the tactics selected for approaching the goal. Within limits of cognitive resources—the student’s cognitive conditions and working memory capacity—and given particular external task conditions—resources, instructional cues, time, and social factors—these updates afford potential for the student to exercise metacognitive control that adapts engagement in mid task.

E. SUMMARY AND PRELUDE TO MEASUREMENTS OF SELF-REGULATED LEARNING

Winne and Hadwin’s (1998) model identifies four phases of a task’s history that are differentiated by products created in each phase. Each product constitutes a different focus for metacognitive monitoring and metacognitive control that enact forms of self-regulation. Although monitoring is common across the phases as the trigger for cognitive control that enacts regulation, information monitored varies as a function of the product in a phase.

We believe this model affords views of SRL that suggest alternative approaches to measuring SRL as an aptitude and as an event. For example, think aloud protocols invite learners to describe unique task conditions and cognitive conditions within which they consider adapting a cognitive tactic based on cognitive evaluations of products created in Phase 3. Survey questions, in contrast, characterize a general or default set of conditions and, in reference to it, invite the student to construct or recollect a memory of some SRL event(s) that is (are) most probable (modal), significant (unique), or representative (median). In the former case, the data available allow measurements of elements in Phases 1 and 2. In the latter case, although elements of Phases 1 and 2 are not part of the observed measurement, they are nonetheless central to how the student responds. These kinds of differences open different windows onto views of SRL. In the next sections, we contrast these measurement approaches using features of Winne and Hadwin’s model.

III. PROTOCOLS FOR MEASURING SELF-REGULATED LEARNING

This section describes seven protocols for measuring SRL. We group protocols according to whether they assess SRL as an aptitude or an event. Due to space limits, we have selected one or two well-known examples of each protocol and described them in detail. These protocols and measures focus on cognition and behavior related to learning in common educational settings such as schools, studying at home, or learning in computer environments.
A. MEASURING SELF-REGULATED LEARNING
AS AN APITUDE

Self-Report Questionnaires

Self-report questionnaires are the most frequently used protocol for measuring SRL, perhaps because they are relatively easy to design, administer, and score. These measures inherently provide (a) information about learners' memories and interpretations of their actions and (b) their explanations of cognitive and metacognitive processes researchers cannot observe (Turner, 1995). Typically, self-report questionnaires measure SRL as an aptitude because items ask respondents to generalize their actions across situations rather than referencing singular and specific learning events while learners experience them.

A survey of conference programs and abstracting services will turn up scores of self-report questionnaires. We describe two of the most used of these, the Learning and Strategies Study Inventory (LASSI; Weinstein, Schulte, & Palmer, 1987) and the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991). Both are frequently used in research and, to our knowledge, are the only two self-report inventories that are accompanied by a manual. Other examples of inventories include the Index of Reading Awareness (Jacobs & Paris, 1987) and the Metacognitive Awareness Inventory (Schraw & Dennison, 1994).

Questionnaires are typically developed according to a three-slot script. First, after items are collected or written, an exploratory factor analysis of a sample's responses is done to examine correspondences between those responses and the model of SRL developers used to generate individual items. Second, reliability coefficients, almost always coefficients of internal consistency, are reported for the full scale and subscales, if there are any. Finally, the full scale and its subscales are correlated with external measures, almost always achievement. We use this script to describe the LASSI and the MSLQ.

Learning and Study Strategies Inventory (LASSI). The LASSI (Weinstein et al., 1987) is a published, standardized, normed 77-item self-report inventory "designed to measure use of learning and study strategies" (Weinstein, 1987, p. 2) by undergraduate students. Items are simple declarations (e.g., I try to interrelate themes in what I am studying) and conditional relations (e.g., When work is difficult, I either give up or study only the easy parts). Students respond using a 5-point scale: not at all typical of me, not very typical of me, somewhat typical of me, fairly typical of me, and very much typical of me. Each of these options is elaborated by a sentence in instructions to students. For instance, "By fairly typical of me, we mean that the statement would be true of you about half the time."
The LASSI User's Manual (Weinstein, 1987) states that "no total score is computed since this is a diagnostic instrument" (p. 3). Uses for scores are suggested. For example, "If a student scores poorly on the Test Strategies Scale, she should be advised to concentrate at least part of her efforts on learning more about how to prepare for and take tests" (p. 3).

The LASSI can be self-administered and self-scored, and a scoring booklet is available for this purpose. It instructs a student how to compute subscale scores by adding items that form the subscale. On a separate page, the student places Xs in a column of numbers labeled by each subscale, then connects the Xs to draw a "learning and study strategies profile." The scores in each column are spaced to correspond to percentiles provided in the table's outside leftmost and rightmost columns. However, no instruction is given for interpreting a percentile other than to indicate it provides a comparison "in relation to other college students answering the same items."

Development of the LASSI included pilot testing of an initial pool of 645 items, interviewing students who took preliminary versions of the inventory, and eliminating items that correlated substantially (≥ .50) with a scale of social desirability. Items on the 1987 version are grouped to form 10 nonoverlapping subscales. On a self-scoring guide, these are labeled (1) attitude and interest, (2) motivation, diligence, self-discipline, and willingness to work hard, (3) use of time management principles for academic tasks, (4) anxiety and worry about school performance, (5) concentration and attention to academic tasks, (6) information processing, acquiring knowledge, and reasoning, (7) selecting main ideas and recognizing important information, (8) use of support techniques and materials, (9) self-testing, reviewing, and preparing for classes, and (10) test strategies and preparing for tests. The inventory has been normed on "a sample of 880 incoming freshman at a large southern university" (Weinstein, 1987, p. 5) and stability coefficients over a 3-week interval were calculated on a sample of 209 students in an introductory communications course. The User's Manual reports coefficients of internal consistency and test–retest coefficients for each subscale. Alphas range from .68 to .86 and stability coefficients range from .72 to .85.

Motivated Strategies for Learning Questionnaire (MSLQ). Pintrich and his colleagues developed the MSLQ (Pintrich et al., 1991) "to assess college students' motivational orientations and their use of different learning strategies for a college course" (p. 3). Items are simple declarations (e.g., I make good use of my study time for this course) and conditional relations (e.g., when reading for this course, I make up questions to help focus my reading). Students record answers to items using a 7-point scale anchored by not at all true of me (1) and very true of me (7). Instructions
are to “answer the questions about how you study in this class as accurately as possible.”

The inventory’s 81 items manifest a hierarchical design, where subscales are nested within sections nested within one of two broad categories, motivation or learning strategies. The motivation category has a value section with three subscales—intrinsic goal orientation, extrinsic goal orientation, and task value—and an expectancy section that consists of three subscales—control of learning beliefs, self-efficacy for learning and performance, and test anxiety. The learning strategies category is divided into two sections. A cognitive and metacognitive strategies section includes subscales labeled rehearsal, elaboration, organization, critical thinking, and metacognitive self-regulation. A resource management strategies section has subscales of time and study environment, effort regulation, peer learning, and help seeking. Subscale scores are means of responses to subscale items after flipping some items that are worded contrary to the subscale’s construct.

The authors provide examples of feedback sheets that can be returned to students. Each sheet provides a short paragraph describing the construct a subscale is intended to reflect; a list of items comprising the subscale; blanks to be filled in for the student’s score, the class mean and scores at quartile boundaries (25th, 50th, and 75th percentiles); and a paragraph of suggestions about how to improve standing on that facet of motivation or learning strategy.

Development of the inventory spanned approximately 3 years, during which time items were tried and revised based on the results of factor analyses, reliability analyses, and correlations with measures of achievement. Because the MSLQ is designed to be used in individual courses, norms have not been developed. The inventory’s manual reports subscale and item-level means, standard deviations, and correlations with final grade for a sample of university (N = 356) and community college (N = 24) students spanning 37 classrooms and 14 subject domains. Alpha internal consistency coefficients for subscales range from .52 to .92.

Separate confirmatory factor analyses have been computed for the motivation section and the learning strategies section using data from the aforementioned sample. The manual (Pintrich et al., 1991) reports that “while goodness of fit indices are not stellar, they are, nevertheless, quite reasonable values, given the fact that we are spanning a broad range of courses and subject domains” (pp. 79–80). We observe that coefficients from this analysis that are parallel to item-factor correlations often are less than a cutoff that the manual describes as appropriate for well-defined constructs. Among the six motivation subscales, phi values (interpreted like correlation coefficients) range from −.17 to .83. Phi coefficients among subscales in the learning strategies section range from .07 to .85. The
authors conclude that "overall, the models show sound structures, and one can reasonably claim factor validity" (Pintrich et al., 1991, p. 80).

Structured Interviews

Interviews span a variety of protocols ranging from a simple query such as, "Tell me about how you ....," to highly structured scripts that list specific questions that are asked verbatim along with rules that control which particular following or follow-up question is posed conditional on information a student has just related. We differentiate interview protocols from a closely related protocol, think aloud (see the beginning of Section III.B), by a simple criterion: If the student is prompted to describe SRL while engaging with an authentic task, the method is a think aloud and SRL is measured as an event. In contrast, if the student is prompted to describe SRL based on memories about what is "typical" of behavior under a certain set of circumstances or to offer judgments about what probably would be typical behavior in a plausible future situation, the protocol is an interview and SRL is measured as an aptitude. Stimulated recall is an interview in which respondents describe their behaviors after completing a specific task and sometimes while reviewing records of engagement with that task, such as a videotape or worksheets. In this case, SRL may be construed either as an event or an aptitude, depending on particulars of the measurement intervention.

Interviews generate verbal descriptions. There are two fundamental alternative approaches to analyzing these descriptions: emergent and theory driven. Analyses that adopt an emergent method seek to identify classes or categories of SRL in the student's descriptions without enforcing a particular or a priori framework. This protocol is "bottom-up" in the sense that features in the material are the origins of classes or categories. We point out that it is not possible to be rid of all one's prior perceptions in applying an emergent method. To do so logically would prevent using any language or learned symbol system to examine, classify, and interpret emergent data. Rather, emergent methods strive to avoid prejudging what should be examined and how information should be classified. The alternative to emergent analyses are theory-guided analyses. These begin with a set of types or classes of data relative to which the students' description of SRL will be classified or quantified. It is a "top-down" approach. These two approaches are readily and often profitably combined in work that tests or extends a theory.

It is a misconception that emergent protocols are devoid of quantitative features. When unique instances of SRL are identified, these instances are assigned a count of 1. When interpretations characterize a class of SRL events as common (or rare), this entails comparing counts, that is, frequencies. If conditional relationships are described, a probability of the transi-
tion from condition to a particular following state is inherent in the description, even if it is not estimated with precision.

**Self-Regulated Learning Interview Schedule (SRLIS).** The SRLIS was developed by Zimmerman and Martinez-Pons (1986, 1988) to explore SRL among high school students. It is a theory-guided, structured interview protocol. Table 2 lists 15 classes of SRL identified by the SRLIS in terms of Winne and Hadwin's (1998) model of SRL (Figure 1). Data about SRL-related behaviors are elicited by having students consider a contextualized but fictitious task. For example, two of six tasks included in the SRLIS are as follows: "Assume a teacher is discussing a topic with your class such as the history of the civil rights movement. He or she says that the class will be tested on the topic. Do you have a method to help you learn and remember what was discussed in class?" and "Most students find it necessary to complete some assignments or prepare themselves for class at home. Do you have any particular methods for improving your study at home?" (Zimmerman & Martinez-Pons, 1988, p. 285). If a student does not answer or indicates the protocol is not understood, a follow-up prompt is given: "What if you are having difficulty? Is there any particular method you use?" (Zimmerman & Martinez-Pons, 1988, p. 285).

<table>
<thead>
<tr>
<th>Class</th>
<th>Description relative to Winne &amp; Hadwin's Model&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-evaluation</td>
<td>Monitoring products relative to standards</td>
</tr>
<tr>
<td>Organizing and transforming</td>
<td>Overt or covert rearrangement of resources</td>
</tr>
<tr>
<td>Goal setting and planning</td>
<td>Adopting or generating standards and planning the sequence, timing, and completion of activities</td>
</tr>
<tr>
<td>Seeking information</td>
<td>Searching for updates to conditions in nonsocial resources</td>
</tr>
<tr>
<td>Keeping records and monitoring</td>
<td>Logging information in an external form about conditions, operations, and products</td>
</tr>
<tr>
<td>Environmental structuring</td>
<td>Selecting or arranging task conditions</td>
</tr>
<tr>
<td>Self-consequences</td>
<td>Updating cognitive conditions (rewards and punishments) following monitoring of a product's standards</td>
</tr>
<tr>
<td>Rehearsing and memorizing</td>
<td>Overt or covert rehearsing</td>
</tr>
<tr>
<td>Seeking help from peers</td>
<td>Updating task conditions by asking peers for information</td>
</tr>
<tr>
<td>Seeking help from teachers</td>
<td>Updating task conditions by asking the teacher for information</td>
</tr>
<tr>
<td>Seeking help from adults</td>
<td>Updating task conditions by asking an adult (e.g., a parent) for information</td>
</tr>
<tr>
<td>Reviewing tests, notes and textbooks</td>
<td>Updating cognitive conditions, such as domain knowledge, and potentially cognitively operating on them again</td>
</tr>
<tr>
<td>Other</td>
<td>Unsolicited task-related information originating with an external source (e.g., peer, teacher, computer)</td>
</tr>
</tbody>
</table>

<sup>a</sup>All are student-initiated except behaviors falling in the Other class.
Three procedures have been used to score interviews (Zimmerman & Martínez-Pons, 1986, 1988). First, a dichotomous score describes whether a student uses a class of SRL. This is based on whether information referring to that class appears at least once in the transcript where the student describes how a task is addressed. Second, a frequency score is created by counting the number of instances of each class of SRL. In their 1986 study, Zimmerman and Pons observed that none of the classes of SRL was mentioned more than once in any of the six tasks, so this score reflected the number of tasks in which a particular class of SRL was used. A third method, yielding what Zimmerman and Martínez-Pons label a consistency score, is to ask students, after they have described how they address a task, to rate how consistently they use the classes of SRL they described. Ratings are made on a 4-point scale: seldom, occasionally, frequently, and most of the time. Although this consistency rating superficially resembles a questionnaire protocol (see the beginning of Section III.A), it differs in the SRLIS in that the student has just generated a description of how a context is addressed. In contrast, in almost all questionnaires, context and tasks within them are quite general such as "when you study" or "in this course."

A significant issue in scoring data generated by interviews is the degree to which scorers agree. After training raters on interview data gathered during pilot work, Zimmerman and Martínez-Pons (1986) reported that two scorers agreed at an 86% level when the number of categories scored identically was divided by the total number of unique categories identified by both scorers.

In the study by Zimmerman and Martínez-Pons (1986), participants attended schools that streamed students into an advanced achievement track or several other lower achievement tracks. The researchers used a dichotomous indication of ability, advanced track or other, as a criterion variable in a discriminant function analysis where predictors were scores on strategy use, strategy frequency scores, and strategy consistency ratings. The analysis was able to identify 91% of students correctly by track with a discriminant function that had standardized coefficients of -.66 for strategy use, .41 for strategy frequency, and 1.12 for strategy consistency rating (all coefficients p < .001).

Teacher Judgments

In their daily interactions with students, teachers are uniquely positioned to judge qualities of students’ SRL. Typically, they base their judgments on event-related measures using trace methods and think aloud protocols that we describe later. In research, however, teacher judgments have been used infrequently as measures of students’ SRL, perhaps because some researchers question the trustworthiness of these data (see Hoge & Butcher, 1984). A particular concern has been teachers' ability to
distinguish between related constructs such as achievement and ability, or achievement and motivation. Reviews, however, indicate that teachers' judgments can avoid these shortcomings and serve as worthy assessments of students' achievement-related behaviors (Hoge & Coladarci, 1989; Perry & Meisels, 1996). According to Perry and Meisels, three conditions are required to achieve this status. First, judgment measures should be direct; that is, ask about distinctive observable student behaviors that reflect a small number of clear and distinct categories or constructs. Second, measures should be criterion-referenced or ask teachers to compare students to a familiar and stable norm group, such as comparing a target student to other students in their current class or to all students they have taught in the past. Third, teachers should be given specific and understandable metrics in terms of which to report judgments. Perry and Meisels also recommend triangulating teacher judgments with data gathered by other protocols.

**Rating Student Self-Regulated Learning Outcomes: A Teacher Scale.** Zimmerman and Martinez-Pons (1988) posited that teachers are uniquely positioned to observe students' SRL during daily classroom activities, and they developed an instrument that meets Perry and Meisels' (1996) criteria. Zimmerman and Martinez-Pons investigated relationships between students' reported use of SRL strategies, teachers' observations of students' SRL in classroom activities, and students' performance on a standardized achievement test, the Metropolitan Achievement Test (MAT). Zimmerman and Martinez-Pons predicted that teachers' ratings of students' SRL would correlate substantially with students' reported use of SRL strategies and moderately with students' performance on the MAT, because standardized measures assess aspects of students' abilities/learning that are distinct from SRL.

In their study, 80 grade 10 students (44 boys, 36 girls) that represented a range of achievement levels first described their SRL in an interview using the SRLIS protocol described previously. When students described one or more strategies for a given task, they were asked to rate the consistency with which they used each one on a 4-point scale.

Three teachers (two males) were asked to rate their own students' SRL as they had observed it in their classrooms using the Rating Student Self-Regulated Learning Outcomes: A Teacher Scale. The 12 items on this protocol were based on the same 14 categories of SRL in the SRLIS (Zimmerman & Martinez-Pons, 1986). Some strategies in the SRLIS are not reflected in the teacher judgment measure because they reflect self-regulating behaviors that typically occur outside the presence of teachers, such as structuring the environment to complete a homework assignment.

Items on the teacher judgment scale were worded as questions about observable behaviors; for example, "Does this student solicit additional
information about the exact nature of forthcoming tests?” Teachers were instructed to rate individual students on each item according to a 5-point scale: 1 = never; 5 = always. Some items referred to students’ direct use of a strategy, such as seeking teacher assistance, seeking information, and self-evaluating. Other strategies, such as organizing and transforming information or goal setting and planning, required teachers to judge observable outcomes of strategy use, such as whether the student completed assignments on or before a specified deadline and whether the student set goals and made plans. The three teachers represented three disciplines—English, social studies, and mathematics—and no two teachers shared students.

The reliability of the teachers’ ratings, according to the Kuder–Richardson formula 20, was .95. An obvious issue in teachers’ judgments of their students’ SRL is whether students’ observable behavior reveals enough of the students’ cognitions. To examine this concern, Zimmerman and Martinez-Pons correlated the teachers’ ratings of students’ SRL with students’ self-reports about using SRL strategies. The correlation was \( r = .70 \). Although this correlation cannot reveal whether teachers’ judgments matched the frequency of students’ SRL, it does show strong proportionality. Consistent with the researchers’ hypotheses, teachers’ ratings correlated moderately with students’ MAT scores, \( r = .43 \).

B. MEASURING SELF-REGULATED LEARNING AS AN EVENT

Think Aloud Measures

The think aloud is a protocol in which a student reports about thoughts and cognitive processes while performing a task. Teachers long have used this method when they request of students, “Explain your work.” The think aloud protocol can be as unstructured as that or it can follow a formal, conditional script that dynamically adjusts which questions or comments an observer makes depending on how the student behaves or whether the student mentions particular information.

A detailed scientific consideration of think aloud protocols and the data they generate was provided by Ericsson and Simon (1984/1993), but the protocol has been used for several thousand years according to Boring (1953). The area in which think aloud approaches to measuring SRL have been most popular is reading. Pressley and Afflerbach (1995) compiled and examined findings from 38 primary think aloud studies dealing with metacognition and SRL in reading. A notable feature of the set of studies they reviewed is diversity: Samples ranged from fourth-grade students to university professors; topics spanned almanacs, poetry, excerpts from short stories and novels, scholarly research publications, and computer programs. Methodologies also varied. Readers were set a variety of goals at
the outset, the two most common being preparing for a test and reading as one normally would to be able to use the information afterward.

In some studies that Pressley and Afflerbach reviewed, the reader's thinking aloud was generated concurrently while reading. There were several variations about when readers were to provide descriptions: whenever the reader wanted (which signals at least that monitoring occurred); retrospectively at points within a task chosen by the observer, as when readers were interrupted every 2 minutes or upon completing each sentence or a section of text; or upon completing an entire reading assignment. Pressley and Afflerbach also noted that the studies they reviewed provided varying instructions about what readers were to report. This feature varied from the observer modeling the procedure of thinking aloud in general, to providing no guidance about what the reader was to report, to naming specific types of events or information (e.g., elaborations, predictions) that were targets of an investigation.

Verbal accounts generated by think alouds are analyzed subsequently using a protocol like that described by Zimmerman and Martinez-Pons (1986). The primary purpose for using think aloud protocols is to map out models of SRL. Except for seeking to determine whether raters agree on assigning readers' descriptions to categories, there is little other standard information about measurement properties of the think aloud protocol.

**Error Detection Tasks**

As noted earlier, metacognitive monitoring that identifies discrepancies between a goal and a current state of a task is the trigger for SRL. Cognitive evaluations produced by metacognitive monitoring then serve as a basis for deciding how to proceed with the task, thus exercising metacognitive control. To measure the monitoring that is a precursor to control, researchers have sometimes introduced errors into materials that students study or use in a task and then observed whether (a) the errors are detected and (b) what students do when they detect them. This protocol is known as the error detection method (Baker & Cerro, in press; Garner, 1987). Several variations of error detection have been developed with respect to (a) the context within which students report errors and (b) measurements that reflect detection (monitoring) of errors and subsequent exercise of metacognitive control.

Regarding contexts for reporting errors, students can be informed about the presence of errors in material after they study it or before. When students are not told about the presence of errors in the material, the context is plausibly more typical of everyday studying, because students justifiably should be able to assume that materials they use in learning have not purposefully been corrupted. In both contexts, it has been shown that a critical feature of measurement is information given to the student about the nature of errors. For instance, students can be given a very
general description of error as “something that might confuse people or something that people might have trouble understanding” (Baker & Zimlin, 1989, p. 342). Alternatively, they can be thoroughly instructed, including being given criterion-referenced tests to validate the effects of that intervention, about types of errors that might appear in material, such as external consistency (created by replacing a word in a sentence with one that makes the meaning implausible or untrue), structural cohesiveness (adding information that fits the general theme of materials, but is irrelevant to the task at hand), and incompleteness (deleting information) (Baker & Zimlin, 1989).

Indicators that students monitor materials for errors have taken several forms including asking students to underline or otherwise mark spots where there is an error. Provided the student limits underlining to these instances, this identifies the extent to which a student discriminates errorful from error-free information. Another kind of data used to indicate monitoring has been eye fixations. This protocol requires an assumption that the duration of a fixation indicates monitoring versus other cognitive operations such as decoding. However, indications of monitoring alone are not sufficient to imply SRL. There also must be an indication of control.

In studies of reading comprehension, students’ tendencies to engage in SRL have proven to be a drawback in measuring students’ ability to detect errors. For instance, in her inaugural study of monitoring while reading a text that was corrupted with errors, Baker (1979) observed that adults who were not informed about the presence of errors appeared quite incompetent in locating them. After the study period, her readers could identify an average of only 38% errors overall and, more importantly, they reported noticing only 25% of the errors while they were studying the material in the first place. On examining recall measures, Baker observed that readers did not misrepresent or omit information, as might have been expected due to errors added to the study material. She interpreted that readers exercised control after recognizing errors by applying automatic fix-up tactics to repair them. In other words, it was because they engaged in SRL while studying that the readers were relatively unable to recall, recreate, or classify errors in the material they studied. In this sense, accurate recall of information in study material that contained an error indicates that SRL was applied provided that reconstructive memory was not the source of correct information appearing in the recall. Moreover, when the student cannot remember that an error was repaired, this plausibly indicates that monitoring and control were automatic, beneath levels of conscious inspection (see Winne, 1995).

**Trace Methodologies**

Traces (Winne, 1982) are observable indicators about cognition that students create as they engage with a task. When a student underlines or
highlights material in a text, for example, this creates a trace signaling that the student discriminated marked content from other information. If the student writes material in the margin that labels or comments about particular marked text, the information provided by that annotation may allow an inference about standards the student used for selecting the marked text. Such annotations also may provide information about metacognitive control exercised. For instance, if the note contained a recognizable mnemonic that was not provided in the text, this reveals the student sought to control the likelihood of subsequent retrieval of information.

To make interpretations about standards used in metacognitive monitoring and cognitive operations applied in exercising metacognitive control, traces must be coupled with a model of cognition. For instance, suppose the student writes “See p. 23” in the margin next to some underlined text on page 30. Under Winne and Hadwin’s (1998) model of SRL (Figure 1), this indicates that one of student’s current goals is to off-load into that permanent comment an otherwise challenging demand on memory to remember that a comparison tactic should be applied to the highlighted information at a later time, say, during review of the material for a test. It might be plausible to infer further qualities about cognitive conditions, such as that the student holds (a) a belief that such demands on memory are likely to exceed capabilities and (b) has an epistemological view that effort applied to link material across pages 23 and 30 has value.

Under current models of memory search (e.g., McKoon & Ratcliff, 1992), the trace just described also implies that the student searched memory for a particular kind of information, namely, comparative information. When one or several candidates were retrieved, monitoring was applied to judge whether its attributes could be compared to information marked by the underline. Because we (presumably) know how the subject matter is organized, it also may be possible to infer standards the student used to monitor the success of the memory search.

Howard-Rose and Winne (1993) described trace methods in a study of SRL. Some traces they observed were unobtrusive. When students underlined, put a check near a block of information (e.g., a cell in a table), or inserted an asterisk into material, they inferred the student had discriminated that marked information from surrounding material. When students copied information into notes from a text or other resources, this was taken to indicate they had rehearsed that information. If the student wrote in their notes a phrase such as, “Think I’ll ask the others in the group what they did for this part,” it was inferred that students held a belief that a social or a material resource could help in addressing the task at hand. Students also were observed to write phrases in notebooks that signaled they were assembling information into larger units, such as when they developed a summary of material or drew a personal analogy to material
they studied. They also tallied instances when students wrote comments about tactical planning, such as "How much of this activity will I need to do?," and monitoring, such as "This task is confusing." (All these examples of student statements were reported in Howard-Rose & Winne, 1993, p. 594.)

One trace in Howard-Rose and Winne's (1993) study was intrusive in the sense that students were instructed to monitor task engagement in ways that were not natural and to record perceptions about that engagement. It concerned attending to thoughts or events not relevant to the task at hand (e.g., I wonder what Joey's doing now?). Students wrote an "A" in the margin of the materials they used when they noticed off-task thoughts or behavior.

Except for concern about observer agreement in coding traces, an issue that is well known and understood, there is little information about measurement issues and uses of trace methods.

**Observations of Performance**

Recently, research on SRL has expanded to investigate relationships between contexts for learning and student behaviors. This reflects contemporary, sociocognitive views of learning that prompt researchers to consider how features that constitute contexts for learning, such as task structures, authority structures, and evaluation practices, influence students' beliefs about themselves as learners, their goals and expectations, and their decisions about how to regulate their behavior in learning activities (Perry, 1998; Pintrich, Marx, & Boyle, 1993).

Observations have three strengths as measurements in research on SRL (Turner, 1995). First, like trace protocols, they reflect what learners do versus what they recall or believe they do. Second, observations allow links between learners' behaviors to task conditions, especially those where feedback is available within the boundaries of a task. Finally, observations can ameliorate difficulties associated with assessing young children's SRL such as positive response bias and the children's limited language for describing cognitive processes.

Turner (1995) observed SRL events to investigate how classroom reading tasks influence young students' use of reading and learning strategies, persistence, and volitional control among average achieving students in first-grade classrooms. The observation system was divided into three sections. The first recorded identifying data: the name of the child being observed, their classroom, and the date and time span of observation. In the second section, observers used narrative protocols to describe students' behaviors and the instructional context as field notes. The third section provided checklists regarding five broad categories of reading strategies that Turner developed inductively by observing students engaged in reading tasks over a 3-month pilot study, and deductively by matching those
records to research on motivation and strategic action. The categories of strategies were decoding (e.g., using graphophonemic or context cues), comprehension (e.g., predicting, summarizing), general learning (e.g., rehearsing, elaborating, organizing), volitional control (e.g., preventing or ignoring distractions), and persistence (e.g., through a difficult task, to correct errors). Observation data were supplemented by student interviews and detailed field notes collected over 5 consecutive days in each participating classroom to characterize teachers' approaches to reading instruction (whole language or basal) and reading tasks/activities in each classroom.

A time sampling procedure was used to collect observation data about SRL. Over the course of a reading or literacy activity, occurrences of any predefined SRL events were marked as present or absent at the end of each successive 3-minute segment. The frequencies of marks over the course of the activity were then dichotomized to reflect the presence or absence of each SRL event within the activity; that is, the measurement protocol was a sign system. Observers sat near target students and recorded in the narrative section frequencies of particular student behaviors as well as verbalizations, facial expressions, gestures such as finger pointing, eye movements, and writing. Immediately following each observed activity, observers interviewed students about the task they had just completed. The interview protocol included four open-ended questions designed to assess students' awareness of two key facets of SRL: the purposes of particular tasks and thinking during task engagement. Two tests of reliability were applied to student observation data: coder speed agreement, which reflected whether observers recorded the same number of behaviors for the same child, and coder category agreement, which addressed the extent to which observers agreed on the category assignment for individual behaviors. Coder speed agreement was 99% and coder category agreement was 89%.

Perry (1998) used Turner's (1995) measurement design as a model for her observational procedures to investigate young children's SRL in select writing contexts. Perry's observation instrument also had three sections. The first section recorded identifying data. The second section was a specimen system (Evertson & Green, 1986), where observers kept a running record of what was going on, including verbatim transcripts of teachers' and students' speech. SRL events were subsequently coded as present or absent.

In Phase 1 of Perry's three-phase study, the third section was a checklist of features of classroom contexts theorized to promote SRL, including items describing students' choices, control over challenge, opportunities for self-evaluation, and teacher and peer support. In Phase 3 of the study, the third section was a checklist of items describing behaviors associated with SRL such as writing strategies, portfolio strategies, executive strategies (e.g., modifying tasks or the environment to control challenge), and
evidence of persistence. After each observation period, events and actions recorded in the running record were matched to items on the checklist. In Phase 1, a rating of 0, 1, or 2 was assigned to items on the checklist to code presence and quality in the observed activity. For example, if students had no choice concerning what to write, the "choice about what" item was rated 0. If student choices were relatively unconstrained, a rating of 2 was recorded. In Phase 3, behaviors associated with SRL were coded present or absent. A second rater was used to test interobserver agreement. Agreement between raters was 88% for Phase 1 observations and 97% for Phase 3 observations.

Turner's (1995) and Perry's (1996) observation systems take first steps toward providing data amenable to examining relationships among instructional contexts, the task conditions in Winne and Hadwin's (1998) model, and forms of SRL observable in students' engagements with tasks. In particular, they illustrate measurement protocols for assessing young children's motivation and competence as self-regulated learners in naturalistic contexts. By combining qualitative and quantitative methods, they also couple descriptions of an individual student's processes with general trends for larger groups (Many, Fyfe, Lewis, & Mitchell, 1996). Finally, they illustrate good practice by triangulating observations with measurements from questionnaires, interviews, and samples of students' work.

IV. ISSUES IN MEASURING SELF-REGULATED LEARNING

Although protocols for measuring SRL are new, principles regarding the validity of measurements are not. By and large, the measures of SRL we reviewed here have been well examined in terms of these principles (see also Pintrich et al., in press). Moreover, in general, each instance of the seven methodologies we reviewed provided models and useful starting points for furthering understandings about SRL as aptitude and as event. Beyond these accomplishments, there remain significant issues to address in striving to advance the quality of measures of SRL and thereby enhance understandings about SRL itself. We consider these issues in five main categories.

A. TARGETS FOR MEASUREMENT

Measurements grow out of a view about what to measure and about a metric for measurements. Along with Pintrich et al. (in press) and Schraw (in press) and Schraw, Wise, and Roos (in press), we note that much basic research remains to develop better models of SRL as guides for developing measures of SRL. However, as we noted earlier, this is a recursive bootstrapping process. Models of empirical phenomena depend on empiri-
cal feedback about their validity. Empirical feedback is gathered by measuring phenomena using protocols that are structured by models as well as measurement issues.

We adopted Winne and Hadwin's (1998) model of SRL as a basis for considering measures of SRL because it synthesizes from the current research literature a wide range of targets that can be measured to achieve a full view of SRL (Figure 1). Some targets—elements of task conditions, performance, and external feedback—are observable in the environment of a self-regulating student. Other targets—elements of cognitive conditions, standards used in monitoring, cognitive representations of stages that describe a task, primitive and acquired operations carried out that constitute engagement with the task, and cognitive evaluations—are constructs that must be inferred on the basis of externally observable behaviors.

We know of no measure of SRL that simultaneously and fully represents all these targets for measurement. Instead, what is practiced and practical is to focus on one or a small set of targets while controlling for others in an immediate periphery and while ignoring yet others outside that periphery that a model labels as extraneous relative to interpretations to be drawn from measurements. For example, a questionnaire protocol typically presumes that students have knowledge of the various study tactics and strategies mentioned in items that ask students to rate a tactic’s or strategy’s frequency of use or importance. In making their judgments about a tactic or strategy, students are directed to use a peripheral context such as “in this course” (Pintrich et al., 1991) that sets some task conditions to a constant value, but excludes others. Extraneous factors are presumed to “zero out” in group data due to randomization, but they remain as complications to interpreting data about an individual.

Boundaries that separate focal targets from peripheral targets from extraneous components are created by one’s model of SRL. Issues of how effectively a measurement intervention operationalizes those boundaries, such that the measurement protocol causes data to be generated that can reflect focal targets validly, are those Cook and Campbell (1979) and Messick (1989) describe in their typologies of factors that affect drawing valid causal inferences. In both systems, the root of these factors is the model of SRL. We take it as obvious that the field must continue to bootstrap models of SRL concomitantly with developing better measures of SRL.

**B. METRICS**

Skirting the mathematical basis of measurement theory (e.g., see Cliff, 1982; Michell, 1990), a metric involves two basic qualities: a unit in terms of which a phenomenon is measured and rules by which measurements can
be manipulated. For instance, Turner's (1995) and Perry's (1998) classroom observation systems for recording students' SRL defined a unit as acts differentiated by breaks in a stream of behavior. Subsequently, acts were categorized.

The operation of categorizing assumes a quality of equivalence among members of the category, specifically, that one instance can be substituted for another with neither loss of information nor introduction of irrelevant variance. This assumption allows a counting rule to be applied to acts within a category that aggregates them unit by unit to form a sum of counts or frequencies. In practice, it is almost always assumed further that a unit in one category is comparable to a unit in another category. This allows frequencies of acts to be compared across categories. It is rare for those who develop protocols for measuring SRL in any of the seven methodologies we reviewed to be this specific in defining measurements. Although such examinations should not consume research efforts, more attention could be given to issues of metrics.

Regarding the unit of a metric, Howard-Rose and Winne (1993) raise questions about the grain size of units. Grain size issues concern the dimensions of a datum that constitutes a unit. For instance, a trace in which a student circles a clause and draws an arrow connecting that circle to a term in a preceding paragraph is a small grain unit relative to a questionnaire item such as “I link up different bits of information.” The labels of tactic and strategy also reflect differences in grain size, the latter being larger grained in two senses: it requires multiple tactics and it involves decision making to select those alternative tactics (Winne, 1995). Another variant of the grain size issue is the time span of a unit. An occurrence event occupies a briefer span of time than a contingency or a patterned contingency. An aptitude is theorized to be enduring, at least over the course of a research investigation that may span a few weeks.

We observe that models of SRL commonly characterize events as small grain units, whereas aptitudes and contextual factors, such as prior knowledge and classroom climate, respectively, are quite large grained. Care needs to be taken when metrics for measurements vary. Now that many macroelements of SRL—planning, monitoring, evaluating—are well documented, researchers are turning attention to more in-depth and fine-grained investigations of individuals' processes as they navigate specific learning activities (Many, et al., 1996; Winne, Hadwin, McNamara, J. K & Chu, 1998a). Assuming normal arithmetic operations are valid within a metric, as is almost always done in research on SRL, researchers need to examine whether usual transformations that standardize units are appropriate for units that originate within metrics of quite different grain size. This issue requires blending basic principles of metrics with other assumptions that underlie statistical treatments of data, and it merits further study given current models of SRL.
Finally, our notion of SRL as event (Winne, 1997; Winne & Stockley, 1998) introduces new issues in measuring SRL. Patterned contingencies that represent dynamic SRL events can be described using statistics adapted from graph theory (Winne et al., 1994), but units can become quite complicated when multiple, fine-grained if–then pieces are aggregated to characterize larger grained SRL strategies. The measurement properties of such large grain patterns are not known. Whereas every model of SRL of which we know includes targets like this, this area needs much work. A first step is to gather extensive descriptions about SRL as events in forms that allow studies of patterned contingencies. Computer-supported learning environments (e.g., see Winne, 1989; Winne & Field, 1998; Winne et al., 1998a) that constitute authentic contexts for studying as they gather trace data about SRL events are promising tools for these efforts.

C. SAMPLING

Every measurement is a sample of behavior. Two fundamental questions should be addressed based on this fact. First, what are characteristics that define the population from which the sample is drawn? Second, what qualities of the population are reflected in the obtained sample, to what degree, and how accurately?

Issues about characterizing the population from which a sample is drawn are evident in measurement protocols where the student's responses are selections of information from memory. Self-report questionnaires, structured interviews, and teacher judgments are three such protocols. To our knowledge, there are no well founded understandings about attributes of the personal “historical” database of events that a student searches as the basis for generating responses to questions.

The same issue applies to measurement protocols such as think aloud, traces, and classroom observations. For example, if a student does not carry out a SRL event, does this reflect (a) absence of critical components for that act of self-regulation, (b) a production deficiency, wherein those components are available to the student but not retrieved and brought into working memory to be topics of metacognitive monitoring and control, or (c) an act of self-regulation, wherein the student exercised metacognitive monitoring and control to omit a potential self-regulating act? Each of these options implies a different population from which measurements are sampled and, thus, has implications that bound valid interpretations about SRL.

A parallel issue concerns contextuality or its opposite, generalizability. To the extent that tasks differ and thereby shape perceptual and other cognitive processing that constitutes SRL, measurements will reflect con-
textuality. While the model portrayed in Figure 1 represents a range of factors that might constitute context, there is little work to date on how such differences affect measurements of SRL beyond documenting that students rate components of SRL differently as a function of task (e.g., Hadwin, Winne, Stockley, Nesbit, & Woszczyna, 1997). We recommend that measurements of SRL include more information than most now do about the products of Phase 1 in self-regulating learning, the definition of the task.

A further issue about individual differences in SRL concerns interactions between domain knowledge, the substance worked on in a task, and SRL per se. Expertise sidesteps some elements of self-regulation because knowledge in the domain of expertise encodes what a novice would need to “directly” self-regulate (Winne, 1995, 1997). We suggest that measurements of SRL should pay more attention to qualities of a student’s domain knowledge to tease apart, as much as possible, SRL events from expertise.

In measuring achievement and some other variables, a distinction can be made between measures of maximum ability and typical ability. A questionnaire item might ask students to rate “how true is it of you” that “I try to play around with ideas of my own related to what I am learning in this course” (item 66, MSLQ, Pintrich et al., 1991). A trace protocol might count “blocks” of handwritten text in the margin of a textbook chapter where the student is judged to do the same. Are these measurements of typical ability or would data differ if students were asked or instructed to do this as often or as well as they could? Measures of SRL should distinguish whether they describe typical or maximum ability, a point Baker and Cerro (in press) also noted about error detection methods in general.

Features of the sample itself vis-à-vis the population also need attention. For instance, when asked to judge categorical features of SRL, such as the presence or absence of various self-regulating acts, how fully is the population of memories searched that are relevant to this question? A particularly interesting issue from our point of view is how the student assigns a value to an instance or description of self-regulation, such as its importance or frequency. What algorithm does the student use: averaging, ordinal dominance reflecting the most frequent or salient category, recency of occurrence, or some other rule? These questions concern calibration, the accuracy with which perceptions correspond to actual SRL events. They all concern decision making because whenever there are options about responses, the student must make a decision about how to match response options to a memory. Beyond Butler and Winne’s (1995) discussion of some paths for exploring these issues, we know of no work on matters such as these that underlie interpretations about how measurement interventions give rise to responses.
D. TECHNICAL ISSUES

The literature on technical qualities of measures is vast and the list of possible issues to be taken up in measuring SRL is correspondingly numerous and novel. To illustrate, if trace measures are measures of ability, new work could be initiated to explore item response characteristics of trace measures as indicators of ability and difficulty. Excluding issues such as these, there remain important technical topics to be given more attention in future work on measuring SRL.

Reliability is a fundamental necessity of high quality measurements. Two methods have been applied regularly in work to date: internal consistency for measures generated by self-report and interobserver or intercoder agreement for the other measurement protocols.

Noticeable by its absence is concern for stability, the extent to which measurements do not vary over time. Stability is a difficult concept to apply to measures of SRL because, by definition, SRL is adaptive and should vary over time under certain conditions. In our view, this tension affords thinking about future work on measuring SRL in two ways. First, high stability should be observed when contexts in which SRL might be applied are homogenous. In contrast, stability should be low when contexts are heterogeneous. This tension may be helpful in bootstrapping better understandings about sampling issues in measurements of SRL. Second, when contexts vary, models of SRL lead to the prediction of high stability for a measure of “adaptiveness.” However, to our knowledge, there are no such measures yet available. One possibility is a measure of conditional probability for transitions to different learning tactics (i.e., applications of metacognitive control) given metacognitive monitoring that identifies differences in context. The graph theoretic measures proposed by Winne et al. (1994) may be a starting point for work on this issue.

Another area requiring attention in measuring SRL links to Cook and Campbell’s (1979) concerns about the construct validity of putative causes and effects, paralleled by Messick’s (1989) focus on construct relevant versus irrelevant variance within a measurement. To date, except for correlations of SRL with achievement, there is little information that can be revealed by multitrait-multimethod investigations of convergent and divergent validity (cf. Zimmerman & Martinez-Pons, 1988). As depicted in Winne and Hadwin’s (1998) model, a variety of cognitive and motivational individual differences are components in an overall portrait of SRL. Multitrait-multimethod studies would help the field focus on the center of SRL and its relationships to peripheral variables. Given that measures of SRL as an event logically entail individual differences regarding beliefs and motivational variables (Pintrich, et al., 1993; Winne & Marx, 1989), it seems imperative that multitrait-multimethod studies receive quick attention.
We note that, except for the LASSI (Weinstein et al., 1987), which provides a manual plus norms based on one university sample, and the MSLQ (Pintrich et al., 1991), which also has a manual and encourages users to create local norms using quartiles, there are very few attempts to standardize and norm measures of SRL. This perhaps reflects several factors: the newness of work on measuring SRL, the field's flexibility in adopting models that guide the development of measurement protocols, and genuine questions about what would be useful norms relative to purposes for measures of SRL. The absence of standardized measurement protocols also may reflect researchers' interests in tailoring measures of SRL to context or students' developmental levels, components explicitly identified in Winne and Hadwin's (1998) model. In light of the list of issues raised so far in this section, standardizing and norming measures of SRL will be a very complex undertaking.

The LASSI (Weinstein et al., 1987) is the only measurement system that invites ipsative (within person) comparisons. Specifically, it instructs each student to draw a profile of scale scores, based on percentile transformations of raw scores relative to the LASSI's norm group, and then to compare levels of scales. Ipsative comparisons are fraught with potential difficulties, two significant ones being the typically very large standard error of an individual's scale score and the implicit suggestion that scores across scales are independent when correlations show they are not. We urge caution in this use of measurements of SRL as well as further research on how improvements might be made in this use of scores describing SRL.

E. UTILITY

Measurements are collected for purposes, and purposes can be served more or less usefully; that is, measurements have varying degrees of utility for particular purposes. The most prevalent purpose of SRL measures is description in basic research. The LASSI (Weinstein, 1987) and MSLQ (Pintrich et al., 1991) invite students to self-diagnose qualities of their approaches to learning and SRL, but formal studies of their diagnostic utility have not been done. We are not aware that any other measures of SRL have been used for formal diagnosis or evaluation in schools or training situations.

For all methods but self-report questionnaires, gathering and scoring measurements of SRL is quite resource intensive. Potential relief may be on the horizon in the form of computer technologies, such as Winne, Hadwin, McNamara, Chu, and Field's (1998) prototype notetaking system, CONOTES. Systems such as CONOTES can administer and score
self-report questionnaires, record traces of SRL as students engage with and adapt during tasks, and use both kinds of measurements as triggers for adapting interaction with the student; for example, by presenting or withholding objectives or self-test questions. Other information to be learned and tools to use in learning, such as a find tool or a note template, also can be made available to a student based on single, profile, or pattern measurements of SRL, thereby extending protocols to adapt dynamically to a student's engagement. This kind of software also could be configured to record audio for collecting think aloud data or to present a structured interview where branches in the interview protocol were followed depending on the student's response to forced-choice items. Work on this front appears to offer considerable promise in measuring SRL as both aptitude and event.

Software systems such as CONOTES will have limited utility in the immediate future for three reasons. First, few researchers or students have access to them. Second, even if such systems are practically unobtrusive to students, the tools they provide as venues for recording SRL events constitute interventions that can alter, sometimes substantially, how students engage with tasks. Third, techniques need to be invented that validly combine information gathered through software protocols with other data gathered through currently popular methods such as portfolio assessments, performance assessments, process approaches to writing instruction, journals, and other self-assessment systems.

Whatever protocols might be practical in educational contexts, two other issues need special attention. First, teachers will need support to learn about and appropriately interpret measures of their students' SRL. Second, developmental differences need investigation with respect to how well measurement protocols work across the age spectrum. This work should address questions about developmental trajectories that bear on the models of SRL that undergird measurement protocols per se.

V. CONCLUSIONS AND FUTURE DIRECTIONS

Research on SRL and measurement protocols used in this research are relatively new and inherently intertwined enterprises. Each helps to bootstrap the other. We adopt the view that a measurement protocol is an intervention in an environment, disturbing it in a fashion that causes data to be generated. Using that data and a logic of causal inference, we infer properties and qualities of a target of measurement. Thus, measurement involves understandings about a target, its environment, and causal rela-
tionships that connect the two. Under this view, measurement is akin to model building and model testing (Cliff, 1982), and, thus, all measures of SRL are reflections of a model of SRL.

We propose that SRL has dual qualities as an aptitude and an event (Winne, 1997; Winne & Stockley, 1998). It is situated within a broad range of environmental plus mental factors and potentials, and manifests itself in recursively applied forms of metacognitive monitoring and metacognitive control that change information over time as learners engage with a task (Winne & Hadwin, 1998). Each of the seven measurement protocols we reviewed—self-report questionnaires, structured interviews, teacher judgments, think aloud measures, error detection tasks, trace methods, and observations of performance—foregrounds different components of conditions, cognitive operations, standards, and event-related change. Matching what a measurement protocol foregrounds to the purpose for measuring SRL takes a first step toward providing a basis for valid interpretations.

In our judgment, three topics about measuring SRL merit special emphasis in future work. First, too little has been achieved yet in measuring SRL as an event. Challenges in this arena are significant. Protocols are needed for collecting longitudinal measurements that span multiple brief episodes, such as activities, as well as extended periods, such as grade levels. In characterizing SRL as an event, point estimates derived from these data, such as means, are not appropriate descriptions. Methods are needed that (a) characterize temporally unfolding patterns of engagement with tasks in terms of the tactics and strategies that constitute SRL and (b) compare patterns over time to reflect regulation per se (Perry, 1998; Winne et al., 1994). In addition, work is needed on how measures of SRL as aptitude and SRL as event can be coordinated to characterize the full spectrum of SRL.

Second, triangulation across measurement protocols is too infrequent. Because each protocol generates slightly different reflections of SRL, a fuller understanding of models and methods can be achieved by using multiple measurement protocols in research. We recognize that measurements generated by different protocols may not be commensurate even if all are indicators of aptitude or of events. The challenge of commensurability affords opportunity to further understandings about technical features of scales, data, rules for aggregating data, and rules for comparing measurements. We predict that advances on these fronts will correspond to advances in modeling SRL.

Third, research into SRL and its measurement has so far included a limited range of populations, most often involving postsecondary students as participants in studies. Very little is known about young children’s SRL. Until measurements are collected across the age spectrum, understandings
about measurement protocols and about developmental trajectories will remain elusive.

ACKNOWLEDGMENT

Work on this chapter was supported in part by grants from the Social Sciences and Humanities Research Council of Canada to Philip H. Winne (No. 410-95-1046, #410-98-0705) and Nancy E. Perry (No. 410-97-1366).

REFERENCES


The self-regulation of behavior involves establishing goals, and monitoring and evaluating behavior. If there is a discrepancy between actual behavior and goals, there is an attempt to modify behavior to eliminate this discrepancy. If goals have been achieved, a person may engage in reinforcing thoughts and/or activities (self-reinforcement).

There may be problems in aspects of the self-regulation of behavior that may contribute to distress. Goals may be too stringent or may be set at levels that an individual may not feel he or she is capable of achieving. A person may engage in perseverative self-monitoring or, at the other extreme, may disengage from or avoid self-monitoring entirely. These and other issues in the self-regulation of behavior are relevant in several areas of clinical psychology including addictive behaviors, coping with health problems, social anxiety, and depression. Strategies and techniques based on a self-regulation model of human behavior can be used for the treatment of clinical disorders and for the management of health problems.

1. INTRODUCTION

Many psychological disorders can be viewed, at least to some extent, as failures in self-regulation. Self-regulation can be defined as the psychological processes that mediate goal-directed behavior in the absence of imme-
This chapter reviews self-regulatory models and their relationships to several representative clinical disorders as well as relevant empirical research, so that the models may be evaluated. Furthermore, the use and potential usefulness of self-regulation in psychotherapy will be discussed.

Most psychological theories of self-regulation are based to some extent on cybernetics, a theory of automatic control systems (Carver & Scheier, 1982; Kanfer, 1975; Watson & Tharp, 1997). A simple example that illustrates the principles of cybernetic theory involves a thermostat found in any home. A thermostat is an automatic device for regulating temperature. It is set to a particular temperature (standard). A thermometer monitors the actual temperature in the home (sensor). A comparator compares the actual reading with the standard. Finally, an activator turns the heater on or off if there is a discrepancy between the actual temperature in the house and the set standard. Thus, the function of the thermostat is to regulate the temperature of the house. If the comparator determines that the temperature is lower than the standard, the activator will perform its function and turn the heater on. Once the temperature has reached the standard, the activator will turn the heater off. The comparator continually monitors the temperature in the house to determine if there is a discrepancy between the actual reading on the thermometer and the standard.

Human behavior can be viewed as being regulated in a similar fashion. We have certain standards or goals, we monitor and evaluate our own behavior to determine if they are meeting the preset goals, and, if there is a discrepancy between our standards and actual behaviors, we attempt to modify our behaviors. In describing the concept of the self, Hilgard (1949) points out that defense mechanisms and self-deception distort the view of the self before outlining two aspects of the self. These are the existence of continuous personal memories and the continual process of self-evaluation (both positive and negative). Self-appraisal forms an important part of the self-regulation of behavior in the normal self. The self, however, is not always involved in the regulatory process. The self-regulation of physiological processes do not rely on the self. For this reason, Block (1996) has argued that the term used should be autoregulation rather than self-regulation. However, in the literature, self-regulation is still the term that is usually used. Thus, self-regulation involves establishing goals, monitoring behaviors, and appraising behaviors to determine if they meet the established goals of the person. A final element in self-regulation is the reinforcement of behavior. Reinforcement may consist of thinking positive thoughts or engaging in pleasurable activities.

A large part of self-regulation theory involves eliminating the discrepancy between actual behavior and goals. Self-discrepancy theory (Higgins, 1987) differentiates between two different types of discrepancies. The first
is between the actual self and the ideal self. Discrepancies of this nature have been shown to result in depression. The second is between the actual self and the ought self. Discrepancies of this nature have been shown to result in anxiety. Carver and Scheier (1999) have pointed out several differences between self-discrepancy theory and their self-regulation of behavior theory. Among these is the lack of explanatory power of self-discrepancy theory with respect to positive affect. Additionally, if there is a discrepancy between actual behavior and ideal or ought self, but there is substantial progress underway to remedy the discrepancy, negative affect should not result as discrepancy theory suggests. Higgins (1997) discusses his theory in terms of regulatory focus. According to Higgins, there are two types of self-regulation. Ideal self-regulation involves a promotion focus (e.g., accomplishment) and ought self-regulation involves a prevention focus (e.g., safety).

Discrepancy reduction is also discussed in other theories. Rogers (1954) discussed psychotherapy in terms of reducing the discrepancy between the actual self and the ideal self. The therapist provides the client with nonjudgemental, genuine, and accepting reflections of the self. A movement toward a more realistic ideal self is made, resulting in greater congruence between the actual self and the ideal self. Discrepancy reduction is also important in motivation theory. Behavior is dependent on both external stimuli and internal factors (Berlyne, 1960). The discrepancy between external information and internal cognitions induces what Berlyne has called intrinsic motivation. The person is motivated to seek ways to reduce this discrepancy. If there is no discrepancy, the external information is neglected. However, interests continually change and the standards against which behavior is evaluated change (Hunt, 1963). Hunt (1961), in discussing organism–environment interaction, notes that if there is not an appropriate “match” between schema (or cognitions) and environmental encounters, this discrepancy (analogous to cognitive dissonance) will motivate the organism to respond so as to minimize this differential.

Several disorders in clinical psychology may be partially due to or may be exacerbated by a malfunction of one or more facets of self-regulation. Individuals may set goals that are too high or that they do not believe they are capable of achieving. They may engage in perseverative self-monitoring and self-evaluation or, at the other extreme, they may cease monitoring their behavior altogether and engage in behaviors that they otherwise would not perform. Finally, individuals may not reinforce positive behaviors.

Self-regulation theory and research has been carried out in many clinical domains. Four areas have been chosen for this chapter. These are the areas in which self-regulation is most relevant. Each of these four clinical areas will be discussed in terms of each of the four aspects of self-regulation (i.e., goal setting, self-monitoring, self-evaluation, and self-
The first section discusses self-regulatory failure and addictive behaviors. The next section also focuses on health psychology and begins by discussing the relationship between the Type A behavior pattern and self-regulation, and leads to a discussion of the use of self-regulatory principles and strategies in the treatment and management of health problems. Next, social anxiety, followed by depression, are each discussed in terms of problems with aspects of the self-regulation of behavior. Finally, implications for treatment are indicated in each section and discussed again as a separate section at the end of the chapter.

II. SELF-REGULATION AND ADDICTIVE BEHAVIORS

Addictive behaviors, ranging from excessive gambling to life-threatening disorders such as bulimia, may be viewed as failures in self-regulation. Opiate addicts have been shown to have more difficulty in self-regulatory functioning than normal controls as assessed by the Scale for Failures in Self-Regulation (Wilson, Passik, Faude, Abrams, & Gordon, 1989). This scale contains 11 items that were applied to participants’ responses to the Thematic Apperception Test (TAT) by blind raters. Among the results, addicts, compared to normal controls, were found to have a poorer understanding of cause and effect relationships, displayed more confusion surrounding temporal mapping, and exhibited a greater lack of planning and a poorer anticipation of the consequences of behavior.

Many causes of failure in self-regulation have been identified. These include a reduction in the monitoring of behavior, difficulty in coping with stressors, and difficulty in focusing attention on the task at hand (Baumeister, Heatherton, & Tice, 1994; Kirschenbaum, 1987). People use various and multiple strategies for coping with stressful situations. These can be classified as task-oriented, emotion-oriented, and avoidance-oriented coping styles (Endler & Parker, 1994). Task-oriented coping (analogous to problem-focused coping) includes strategies aimed at solving the problem. Emotion-oriented coping includes strategies such as self-preoccupation or daydreaming. Avoidance-oriented coping can include responses that are a distraction from the stressful situation or a social diversion. Some strategies are more effective than others depending on the situation. However, in general, a strong positive correlation is often found between the use of emotion-oriented coping strategies and measures of psychological distress such as state anxiety. There are also individual differences in coping with stress (Endler, 1988, 1997).

Control plays an important role in coping with stress. A recent experiment in our lab investigated the effects of perceived control on psychological distress associated with a contrived stressor (an anagram task) (Endler, Speer, Johnson, & Flett, 1998). Participants with higher perceptions of
control over the situation were found to use more task-oriented coping and less emotion-oriented coping, and experienced lower state anxiety (the measure of psychological distress). Thus, the perception of having control over a situation results in less psychological distress. This study also investigated the effect of experimentally produced control on levels of state anxiety and coping strategies. No effect was found on these variables. Therefore, whether or not one actually has control in their environment is not the important factor; rather, the perception of control plays an important role in coping.

Because having control over a situation usually results in less psychological distress, it would seem that control would be an important issue for individuals with addictive behaviors. Again, it is the perception of control rather than objective control that matters. Because perceptual control is associated with less distress, it is adaptive. Individuals with high perceptions of control would be more likely to effectively cope with stressful situations and hence there would be a lower probability of self-regulatory failure.

For a person to effectively self-regulate their behavior, higher order processes must override lower order processes (Baumeister et al., Carver & Scheier, 1999). Higher order processes involve more abstract long-term goals and lower order processes involve more concrete goals. For example, if an individual is addicted to alcohol, but has decided that it is time to quit that habit, the desire to have a drink would be the lower order process and the decision to stop drinking would be the higher order process. If the individual is able to overcome the desire and abstain from drinking, then the individual has effectively self-regulated his or her behavior. If the person has a drink, a failure in self-regulation has occurred. This situation also can be analyzed in terms of a self-regulatory feedback loop. The desire to stop drinking would constitute the goal. The individual would continuously monitor and appraise his or her behavior to determine if there is a discrepancy between the goal and the behavior. However, there may be a problem in each of these areas.

A. GOAL SETTING AND ADDICTIVE BEHAVIORS

A person may not establish goals or may have conflicting goals (Baumeister et al., 1994; Karoly, 1993). Confusion would result, leading to difficulty in the monitoring of behavior. The individual with the goal of not drinking alcohol also may have a conflicting goal to go out and socialize with friends on a more regular basis. The individual's friends may drink in social situations and may encourage the habit of drinking in social situations. Thus, this person is placed in a high risk and conflicting situation with respect to the original goal when attempting to achieve the conflicting goal. It may be wise to avoid socializing with these friends so as to achieve
the goal of abstaining from alcohol, but, on the other hand, this would contradict the goal of increased socialization. These conflicting goals cause confusion for the individual and may result in a failure in self-regulation.

Another problem related to goal setting is the failure to disengage when a goal is unattainable (Carver & Scheier, 1999). Distress will result from the continued attempt to attain a goal that is unattainable. Disengagement may involve reducing the goal to a less demanding one or abandoning the goal altogether.

Future research should examine the goal setting behaviors of various populations of addictive behaviors. Furthermore, a comparison of goal setting behaviors relevant to the addictive behavior and goal setting in other areas should be made.

B. SELF-MONITORING AND ADDICTIVE BEHAVIORS

Second, a person may stop monitoring his or her behavior (Baumeister et al., 1994; Kirschenbaum, 1987). This is a central cause for failure in self-regulation. The individual loses self-awareness and engages in activities that he or she otherwise would not engage in. Because the person is not monitoring his or her behavior, it also is not evaluated. The person with the desire to stop drinking may binge while in such a phase.

Schupak-Neuberg and Nemeroff (1993) found evidence for this type of failure in self-regulation in a study investigating the notion of sense of self among a bulimic population. Their study compared bulimics and binge eaters on a scale (The Binging Inventory) designed by the authors to assess the extent to which individuals experience a lack of self-awareness during a binge. A binge eater engages in episodes (binges) in which enormous quantities of food are ingested. In addition to binging, a bulimic engages in purging behaviors that include vomiting or the use of laxatives or diuretics. A binge was found to be an escape from self-awareness for the bulimic group, but not for the binge eaters. Furthermore, the purpose of purging was assessed in the bulimic group and it was found to relieve negative affect. Together the binge and the purge cycle was found to serve an emotional regulatory function. Finally, compared to normals and binge eaters, bulimics were found to have greater confusion regarding sense of self. Thus, according to this study, during a binge bulimics are not monitoring and appraising their behavior, so they are, in the moment, unaware of the discrepancy between their goal (which would presumably be to achieve healthy eating habits) and their behavior (i.e., binging).

Other relevant evidence comes from the weight reduction literature. Individuals who engaged in self-monitoring were found to lose more weight than individuals who did not (Bellack, Rozensky, & Schwartz, 1974). Furthermore, it was found that self-monitoring is essential for continued weight loss after the termination of a treatment group (Perri et al.,
McAllister, 1988). Without the continued support and attention of the treatment group, the dieter may not recognize behavior that is not in accordance with his or her goal if he or she is not continually self-monitoring eating and exercise habits. Self-monitoring enhances the maintenance of weight loss after treatment, partially because of the increase in awareness of weight gain and potential relapse (Westover & Lanyon, 1990). Otherwise weight gain may go unnoticed to the point where the dieter may feel that it is too late to continue with the program. Self-monitoring is currently included in most behavioral weight loss programs (Johnson & Baggess, 1993).

It is apparent that the cessation of the monitoring of behavior is implicated in some addictive behaviors. It is unclear whether it plays a role in the onset or only in the maintenance of the addictive behavior. Future research needs to be conducted to make this clarification.

C. SELF-EVALUATION AND ADDICTIVE BEHAVIORS

Disengagement from the monitoring of behavior was shown to be implicated in some addictive behavior populations. Individuals who are not monitoring their behavior are not able to evaluate their behavior. Additionally, a person may not make the necessary changes in behavior if there is a discrepancy between his or her behavior and his or her goal. In this situation, the person sets a goal, and monitors and appraises his or her behavior, but if the behavior falls short of the goal, the individual does not react appropriately. The person may feel unable to effectively change his or her behavior, so the individual does not make any attempt to bring the behavior closer to personal goals. The person who wants to quit drinking will know that is the goal, and will monitor and appraise his or her behavior. This person will be aware that having a drink is contradicting his or her goal to quit drinking. However, when this person does have a drink, he or she will not try to eliminate the discrepancy between behavior and goals (i.e., stop drinking). Persons may feel that they do not possess the strength to accomplish this change in their behaviors.

Perceived self-efficacy, the belief that one is capable of carrying out a desired behavior (Bandura, 1977, 1997), is an important factor in self-regulation. It has been theorized that drug and alcohol addicts use substances as a means of self-regulation (Khantzian, 1990). These individuals become accustomed to using a substance for the purposes of calming themselves or relating to others, for example, and then underestimate their own ability to carry out these functions. They believe that they will not be able to achieve their goals without the use of the substance, which leads to continued substance abuse. Furthermore, the confidence in one's ability to abstain from the addictive behavior and the confidence in one's ability
to recover if there is a setback are critical in the change process (Marlatt, Baer, & Quigley, 1995).

Empirical research needs to be done in this area to determine if individuals with various addictive behaviors are in fact not monitoring or evaluating their behaviors. Furthermore, it needs to be determined if individuals who are monitoring and evaluating their behaviors are concluding that they are unable to minimize the discrepancies between their desired goals and their actual behaviors.

D. SELF-REINFORCEMENT AND ADDICTIVE BEHAVIORS

The amount of self-reinforcement that an individual with an addictive behavior engages in has been theorized to play a role in the onset or the maintenance of the addictive behavior (Miller, 1987). Studies have been conducted to investigate the self-reinforcement behaviors in substance abuse and eating disorder populations. In an undergraduate population, participants indicated that they were not likely to use strategies related to self-reinforcement to control the amount of alcohol consumption. Furthermore, substance abuse patients were found to have reinforcement practices that were no different from other psychiatric groups or normal controls (Parmar & Cernovsky, 1993). Additionally, Rozensky and Honor (1984) compared the self-reinforcement and self-punishment behaviors of psychiatric patients, alcoholic patients, and nonclinical controls. These groups did not differ on self-reward; however, the alcoholic group was found to be the least self-punishing. It appears as though self-reinforcement practices of substance abuse individuals are no different than normals. Within an addiction rehabilitation population, reinforcement scores were found to be related negatively to depression (Cernovsky, 1989). Thus, frequency of self-reinforcement may have a mediating effect. Finally, therapy for bulimia, which involves modifying the self-reinforcement practices of the individual, has been shown to be effective (Mizes & Lohr, 1983).

The relationship of self-reinforcement behaviors to an addictive behavior may be dependent on which addictive behavior is in question. Therapy aimed at increasing the self-reward practices of an individual may be effective with bulimic individuals, but not with substance abuse individuals. Further research should investigate practices of self-reinforcement of various populations of individuals with addictive behaviors as well as determine if therapy aimed at increasing positive self-reinforcement is effective. It may be the case that self-reinforcement is at a similar level to normal controls; however, therapy aimed at increasing self-reinforcement may still prove to be effective.
E. IMPLICATIONS FOR TREATMENT

Self-regulation plays an important role in human adaptation to life. Rothbaum, Weisz, and Snyder (1982) outlined two processes that form adaptation. These are primary control, changing the environment to suit the self, and secondary control, changing the self to suit the environment. Self-regulation is a large component of secondary control. Having the capacity to effectively regulate one's behavior is important for becoming a well-adjusted person and adapting to the environment (Baumeister et al., 1994). Thus, learning to effectively self-regulate behavior is important for therapy.

Raising a client's belief in their self-regulatory efficacy may have an impact on their behavior (Bandura, 1997). In the eating disorder literature, this finding has been observed with bulimic participants (Schneider, O'Leary, & Agras, 1987). The greater their belief in their capability to effectively regulate their behavior, the less they engaged in purging behaviors.

In treating addictive behaviors, Kirschenbaum (1987) discussed "obsessive-compulsive self-regulation" for the prevention of self-regulatory breakdown. The term "obsessive-compulsive" traditionally is used to refer to individuals who have a disorder in which they have unwanted repetitive thoughts, images, or impulses (obsessions), and in response to these thoughts engage in repetitive, unwanted actions (compulsions). Here, the term refers to the constant and detailed self-monitoring of behavior. The intensity with which individuals with obsessive-compulsive disorder are involved with their maladaptive behavior is analogous to the intensity with which persons trying to avoid self-regulatory failure should monitor their behavior. In treating addictive behaviors, it is very important for the client to self-monitor and self-evaluate behavior so as to be aware of discrepancies between goals and actual behavior.

Relapse is an important issue in addiction. Alcohol dependence has a high rate of relapse after treatment (Litman, Eiser, & Taylor, 1979). Often, an initial lapse will lead to a complete breakdown in self-regulation and consequently relapse. An initial lapse may occur while the person is in a high risk situation. The individual may be experiencing a considerable amount of negative affect due to a variety of sources. There may be pressure from peers to engage in the unwanted habit. Once the initial lapse has occurred, it is important that the individual does not relinquish all efforts to succeed at his or her goal. The person may feel guilty about the lapse and feel that there is nothing he or she can do about it (Curry, Marlatt, & Gordon, 1987). If self-monitoring and self-evaluation continue to take place, the individual may be able to get back on track. However, the danger of the cessation of self-monitoring is paramount once a lapse has occurred. A study compared adults who were in a smoking cessation
treatment group who lapsed, but then stopped smoking once again, and adults from the same group who also lapsed, but for which the lapse lead to relapse. They found that a plan to cope with an initial lapse was the important factor (Candiotte & Lichtenstein, 1981). All of the individuals who were able to avoid relapse had a plan to cope, whereas only half of the relapse group had such a plan.

Related to relapse, timing is also an important factor in self-regulation with respect to some addictive behaviors. For example, if a person is trying to diet, it is easier to refuse a bag of chips right from the original offer. It would be even easier if the dieter avoided situations where unhealthy foods were available. However, if the dieter has one chip, it will be harder and will become increasingly difficult with each chip eaten. Baumeister et al. (1994) label this factor in self-regulatory failure as psychological inertia.

The disengagement of self-monitoring and hence self-evaluation appears to be an important aspect for addictive behaviors. The impact of the role of goal setting and self-reinforcement is not as clear. Other factors such as withdrawal for substance abusers and social factors also have an impact in treatment.

III. SELF-REGULATION AND HEALTH

Self-regulatory principles have been used in the treatment or management of various health problems (in addition to the treatment of addictive behaviors discussed in the previous section). Therapies based on self-regulatory principles have been developed for use with patients with hypertension (Linden, 1988), asthma (Creer & Winder, 1986), diabetes (Wing, 1992), and chronic pain (Grunau & Craig, 1988). Additionally, such therapies have been effective with coronary heart disease patients (Clark, Janz, Dodge, & Sharpe, 1992) and for reducing Type A behavior (Suinn, 1982).

This section begins by outlining the various strategies that patients use for coping with health problems, followed by a description of the Type A behavior pattern. As in section II, each aspect of self-regulation then will be addressed separately.

As previously stated, difficulty in coping with stressors is one of the causes for failure in self-regulation (Baumeister et al., 1994; Kirschenbaum, 1987). Coping strategies for stressful situations already have been outlined. Other research has investigated specific strategies for coping with health problems. Such research has used populations of individuals with both acute health problems (e.g., colds and flus) and chronic health problems (e.g., cancer and cardiac patients). Patients use various and multiple strategies for coping with health problems. These can be classified as distraction, palliative, instrumental, and emotional preoccupation coping strategies (Endler, Parker, & Summerfeldt, 1993, 1998). Distraction
strategies include thinking about other experiences, spending time with others, and engaging in activities unrelated to the illness. Palliative coping strategies are aimed at alleviating the unpleasantness of the illness and include soothing, self-care behaviors such as getting rest. Instrumental strategies are task-oriented efforts such as taking medication and actively seeking medical help. Emotional preoccupation strategies are emotion oriented and include ruminating about the health problem. As with coping with stress in general, different strategies are associated with different outcomes. Instrumental strategies are negatively correlated with psychological distress and length of hospital stay, whereas researchers have found a positive association between the use of emotion-oriented strategies and psychological distress and length of hospital stay (for reviews, see Endler & Parker, 1994, 1999). Effective coping is important to decrease the probability of failure in self-regulation.

Before looking closely at each facet of self-regulation, the Type A behavior pattern and its relationship to health will be described. Self-regulation processes have been theorized to underlie Type A behavior. Defining characteristics of this behavior pattern include heightened ambition, competitiveness, easily evoked hostility, and a rushing or hurrying component in everything that is done. Healthy individuals who exhibit Type A behavior are twice as likely to develop coronary heart disease compared to healthy Type B individuals. The hostility component of the Type A behavior pattern has been found to be the most important aspect with respect to heart disease (Matthews, 1988). The increase in the probability of heart disease associated with the Type A personality is presumably due to increased physiological arousal (Houston, 1983). Various health problems will be discussed in the following sections on the different aspects of self-regulation. Additionally, because of the association between the Type A personality pattern and coronary heart disease, Type A behavior also will be included in each of these sections. Although goal setting, self-monitoring, self-evaluation, and self-reinforcement are areas on which some treatment programs for health problems have focused, there have not been many well controlled studies to determine the efficacy of these approaches. Each section will begin with a review of the research in Type A behavior and then move on to discuss the relevance of that particular aspect of self-regulation with the management of health problems.

A. GOAL SETTING AND HEALTH

Rather than act in ways to reduce increased physiological arousal, Type A persons seem to behave in ways that maintain it. They may associate this aroused state with achievement. Type As have been found to set higher goals for themselves and are more likely to expect to attain their goals than Type Bs (Grimm & Yarnold, 1984; O'Keefe & Smith, 1988).
The use of goal setting in the management of health problems may result in better adherence to the treatment protocol and better outcome. Goal setting has been found to be effective in improving regimen adherence for diabetic patients (Delamater et al., 1991). Patients who set goals for their treatment are more likely to comply with the treatment regimen. The setting of goals may increase the patients' awareness of the desired outcome. Patients are not always clear on their illness and the recommended treatments. Setting goals may help clarify what they should do to manage their health problem most effectively. The practice of goal setting also may have its effect by increasing the patient's attention to their treatment. Because they have a goal to attain, more time may be spent monitoring and evaluating their behaviors relevant to their health problem. The efficacy of goal setting was supported in a diabetic population (Delamater et al., 1991). Future research should look into the relevance of goal setting in the management of other health problems.

Not only is the practice of goal setting important, but the level of goals that are set is an important variable with respect to health problems. Emmons (1992) examined goal setting behavior and divided participants into high-level striving and low-level striving. It was found that high-level striving was associated with higher levels of psychological distress, whereas, low-level striving was associated with higher levels of physical illness.

B. SELF-MONITORING AND HEALTH

Type A individuals tend to focus on negative feedback, which contributes to their behavior pattern and may be maladaptive (Cooney & Zeichner, 1985). It appears as if they are monitoring their behavior, but they are ruminating about negative evaluations. Rumination is not a healthy adaptive response and is associated with distress (Nolen-Hoeksema, 1991).

As for the management of health problems, setting goals may increase the likelihood of self-monitoring. Vigorous self-monitoring is an important aspect for individuals with some health problems. Treatment for type 2 diabetes can involve the careful monitoring of blood sugar levels through diet and exercise programs (Wing, 1992). Similarly, management for coronary heart disease typically also involves careful monitoring of diet and exercise. Monitoring can be important in the management of health problems in which the individual has control over some aspect of his or her illness.

C. SELF-EVALUATION AND HEALTH

Individuals with the Type A behavior pattern are less satisfied with and more critical of their performance (O'Keefe & Smith, 1988; Ovcharchyn, Johnson, & Petzel, 1981). It appears as though Type A individuals monitor
and evaluate their behavior, but they set higher goals and then ruminate about discrepancies between actual behavior and goals.

Self-evaluation can affect physical and psychological health. Individuals who appraised themselves as ineffective problem solvers were found to report more psychological distress (Heppner, Kampa, & Brunning, 1987). Evaluating behavior is an important step in self-regulation. If a diabetic is supposed to monitor his or her diet, he or she must evaluate simultaneously whether his or her food intake is consistent with the treatment protocol. If there is a discrepancy between the actual food intake and the desired food intake, a change in eating habits should take place. Thus, it is important to make a change in behaviors upon the discovery of discrepancies between actual behaviors and goals. Unfortunately, in the same study (Heppner et al., 1987), individuals who appraised themselves as effective problem solvers were found to be higher on the Type A behavior pattern.

D. SELF-REINFORCEMENT AND HEALTH

Compliance with long-term medical regimens is problematic in that only half of all patients with chronic diseases adhere to their treatment protocol (Epstein & Cluss, 1982). Reinforcement for taking medication has been shown to be effective in increasing compliance (Epstein & Cluss, 1982). Compliance also was found to increase in diabetic individuals who engaged in self-reinforcement for adhering to diet and exercise programs (Heiby, Gafarian, & McCann, 1989).

Finally, Type A individuals have been found to be less self-reinforcing (Holden & Wagner, 1990). Reinforcement may involve an activity that can be engaged in or simply may be thinking positive thoughts. Gender differences have been found in this area of research. Compared to Type A females, and Type A and B males, Type B females have been found to be the most reinforcing and the least self-punishing (Holden & Wagner, 1990).

Thus, overall, Type A individuals set higher goals (which they expect to attain), are more critical of their performance, and are less likely to positively reinforce their behavior than Type Bs. The aggressive behavior of Type A persons may be viewed as an attempt to reduce the discrepancy between their preset goals and their actual performance. Implications for the treatment of Type A behavior pattern individuals will be discussed in the next section.

E. IMPLICATIONS FOR TREATMENT

Therapies based on principles from self-regulation theory have been developed for use in clients with various health problems including hypertension, asthma, and chronic pain. The appeal of self-regulation therapies for the management of health problems is that they enable patients to do
something for themselves. Self-regulation therapy may be used in combination with other forms of treatment (i.e., drug therapy) or on its own. This is an especially important form of treatment for patients who do not respond to medical treatment. Most self-regulation therapies involve an effort to increase the client's awareness of behavior through self-monitoring. Relaxation techniques and biofeedback may be used to help clients become aware of different states (Linden, 1988; Shabsin & Whitehead, 1988). Other techniques involve helping the client learn effective coping strategies.

Burke (1989) outlined four basic self-regulatory tools that can be used alone or in behavior therapy to counter the stress response. As previously stated, the inability to deal with stress is a common cause for self-regulatory failure. Thus techniques to help cope with stress are very important. Cannon (1932) identified the stress response as an activation of the sympathetic nervous system that results in many bodily changes including changes in blood pressure, heart rate, and digestion (i.e., "fight or flight" response). Selye (1956) first demonstrated that constant activation of the sympathetic nervous system can lead to physiological damage and sometimes this damage can be irreversible. Tools to prevent or decrease the stress response are especially important in some health problems. For example, constant elevation in blood pressure can lead to hypertension. These self-regulatory tools are meditation, hypnosis, progressive relaxation, and biofeedback.

Meditation involves a passive attitude, a quiet environment, a body position that is associated with low muscle tension, and reduced awareness of the environment achieved through various methods including fixating on an object. Studies have shown that meditation results in relaxation (for a review, see Benson, 1975). These same elements are true of hypnosis. Additionally, in hypnosis, a practitioner typically influences the client during the state of relaxation. The efficacy of hypnosis has been under much debate. Progressive relaxation techniques also involve the elements of meditation. Additionally, the client is instructed to tense and relax different muscle groups in succession. The awareness of tension and relaxation are important for these self-regulatory tools.

Biofeedback is a method of increasing a person's awareness to smaller changes in biological functioning than normally would be noticed. The brain receives information regarding problems in the organs of the body from a feedback loop so that it can regulate the system. Normally, we are not aware of this process until certain levels of distress are reached. Physical or psychological symptoms may appear at this point. Brain wave activity (through the use of an EEG), muscle activity (through the use of an EMG), skin temperature, and sweating (GSR) are among the functions that are monitored using biofeedback. The patient is provided information regarding one of these physiological functions through auditory or visual
stimuli. The client learns to change the intensity of the stimulus and thereby learns to change the biological process. The client can then monitor the process individually and make appropriate changes if necessary. Biofeedback is used for relaxation purposes and can be used in combination with progressive relaxation techniques. All of these tools can be important in decreasing stress and in combating failure in self-regulation.

Biofeedback and relaxation techniques are used to treat various health problems. Outcome studies have shown efficacy for these techniques with hypertension patients (Linden, 1988) and gastrointestinal tract disorders (Shabsin & Whitehead, 1988). Research into the use of biofeedback with pain patients has produced varied results, depending on the type of pain (e.g., headache, migraine, or back pain) and type of feedback (e.g., muscle contraction or temperature) Relaxation and coping skills training are showing potential in the treatment of pain (for a review, see Grunau & Craig, 1988). These tools also are used in therapy with Type A individuals.

Because of the association between the hostility component of Type A behavior and coronary heart disease, it is important to treat this Type A behavior pattern. However, Type A individuals have associated their behavior with achievement, so they may be reluctant to change. An unwillingness to modify behavior especially may be present in Type A individuals who do not have any symptoms of coronary heart disease. Unfortunately, sometimes symptoms of heart disease (i.e., chest pain, fatigue) go unnoticed or are mislabeled by Type A individuals (Frautschi & Chesney, 1988). Lab research has shown that Type As suppress symptoms (Carver, Coleman, & Glass, 1976) and Type As have been found to present with symptoms of myocardial infarction at a later point in their cardiac illness than Type Bs (Matthews, Siegel, Kuller, Thompson, & Varlat, 1983).

Cognitive behavior therapy (CBT) has been shown to be effective in decreasing Type A behavior (Suinn, 1982). A main component in this therapy involves training the client to self-monitor his or her behavior. Type A individuals need to pay more attention to their arousal, thoughts, and behaviors. Relaxation and biofeedback training can be used to make the client more aware of various symptoms of arousal (Suinn, 1982). The use of self-reinforcement of positive changes in behavior and thoughts may help. Type A behavior is reinforced by society and Type A individuals reinforce themselves for their achievements, which are associated with their states of arousal. The client has to recognize this pattern of conditioning that has occurred in the past and also has to be aware that it is unhealthy. Through self-monitoring and self-evaluation, Type As can become aware of discrepancies between their behaviors and their goals (i.e., less Type A behavior), and work to eliminate them. Furthermore, reinforcement of positive changes in behavior can help to reverse the previous patterns of reinforcement.
A program designed to increase patients' knowledge of the processes of self-regulation was shown to be effective for older adults with organic heart disease. Participants in this program were found to do better on measures of psychosocial functioning and had less severe symptoms (Clark, et al., 1992).

Self-management therapy programs have been developed for asthma patients (Creer & Winder, 1986; Hindi-Alexander & Cropp, 1984). Clients are first taught about their illness (e.g., breathing mechanism) and then are taught self-management skills. A main component is self-monitoring. Clients have to monitor their breathing and their asthma condition, and evaluate whether there is a problem. If there is a problem, they must come up with a solution (i.e., eliminate the discrepancy between their goal of being healthy and their current condition). Studies have shown that this is a practical method of controlling asthma, especially for children (for a review, see Creer, 1988). Asthmatic children otherwise may not recognize potentially life-threatening situations.

Thus therapy for health problems can involve learning effective coping strategies to decrease distress and subsequently decrease failure in self-regulation. In Type A behavior pattern individuals, specific aspects in the self-regulation of behavior may be addressed. Finally, in the management of health problems, specific self-regulatory tools (e.g., biofeedback) can be used.

IV. SELF-REGULATION AND SOCIAL ANXIETY

This section discusses aspects of the self-regulation of behavior that may contribute to social anxiety, one of the facets of anxiety. Anxiety has a state and a trait component. State anxiety is a transitional and emotional condition, whereas trait anxiety is a relatively stable personality characteristic (Endler & Magnusson, 1976). Spielberger's (1983) measure of trait anxiety assumes that anxiety is unidimensional and focuses on anxiety in interpersonal situations. The multidimensional interaction model states that trait (as well as state) anxiety is multidimensional and social-evaluation anxiety is one of the dimensions (Endler, 1983, 1997; Endler, Edwards, Vitelli, & Parker, 1989). Social-evaluation anxiety can be defined as anxiety that arises in situations where an individual is being observed or evaluated by others. Other labels for social anxiety are often used in the literature. Rosenberg's (1965) "evaluation apprehension" falls under the category of social-evaluation anxiety. Watson and Friend (1969) defined (and assessed) social anxiety as a combination of the experience of distress in social situations, the avoidance of social situations, and the fear of negative evaluations from others.
Self-regulation models of social anxiety state that social anxiety arises if an individual has not behaved in a manner consistent with his or her original goal (Carver & Scheier, 1986; Schlenker & Leary, 1982). Schlenker and Leary provided a self-presentation model of social anxiety. Social anxiety arises when an individual wants to make a certain impression on others, but feels that success will not occur. The amount of anxiety experienced depends on how close the person believes that he or she will come to achieving his or her goal, the reactions of others, and the importance of the interaction. Individuals assess their behavior and this assessment begins if they consider the interaction important or if they feel they are not achieving their goals. Carver and Scheier (1986) stated that human behavior is controlled by feedback loops. People have a "reference value" or anchor point for how they want to behave and they compare their behaviors to this reference. If their behavior is not acceptable, then a change occurs to make the behavior closer to the reference value or standard. This assessment and appraisal process continues. These basic self-regulatory principles have been applied to social anxiety (Carver & Scheier, 1986). The four components of self-regulation (i.e., goal setting, self-monitoring, self-appraisal, and self-reinforcement) will be discussed separately with respect to social anxiety.

A. GOAL SETTING AND SOCIAL ANXIETY

There has not been much empirical research on goal setting in socially anxious individuals, although it is part of some theories of social anxiety (Carver & Scheier, 1986; Arkin, Lake, & Baumgardner, 1986). Some people have argued that socially anxious individuals believe that others have high standards for them to meet (Rehm, 1977; Schlenker & Leary, 1982). However, Wallace and Alden (1991) found that anxious subjects did not set higher goals for themselves than nonanxious subjects and did not rate others' standards as higher than did the nonanxious group. Although it has been found that socially anxious subjects do not objectively rate others as having higher standards for them, they do seem to believe that others have high standards for them (Alden, Bieling, & Wallace, 1994; Wallace & Alden, 1991). The Socially Prescribed subscale of the Multidimensional Perfectionism Scale was used to measure the standards that the subjects believe others have for them (Hewitt & Flett, 1991). It seems that socially anxious subjects believe that others have perfectionistic standards for them, but this is not demonstrated in their objective ratings. Doerfler and Aron (1995) also found that socially anxious and normal participants did not differ in their goal setting, but the socially anxious participants, however, did not expect to achieve their goals. Thus it may be that socially anxious individuals set goals that are comparable to nonanxious people,
but that they do not expect to attain their goals. They believe they lack the social efficacy necessary to meet the standards of others (Bandura, 1997).

B. SELF-MONITORING AND SOCIAL ANXIETY

Socially anxious individuals may engage in perseverative self-monitoring and self-appraisal, which may contribute to their anxiety. The tendency to be aware of oneself as a social object is measured by the Public Self-Consciousness subscale of the Self-Consciousness Scale (Fenigstein, Scheier, & Buss, 1975). Several researchers have found a correlation between public self-consciousness and social anxiety (Buss, 1980; Fenigstein et al., 1975; Lennox, 1984; Pilkonis, 1977). However, some researchers have not (Linder & Der-Karabetian, 1986). Monfries and Kafer (1993) found that the two constructs are correlated, but when the social distress and social avoidance components of the Social Avoidance and Distress (SAD) scale were separated, public self-consciousness correlated only with the social distress component. This research has relied solely on self-report. Future research should use other methods to provide confirmatory evidence of the existence of this relationship. Additionally, research should be conducted to determine if perseverative self-monitoring and self-evaluation are precursors to social anxiety, if they are factors in the maintenance of social anxiety, or if they play a role both in onset and maintenance.

C. SELF-EVALUATION AND SOCIAL ANXIETY

Socially anxious individuals may evaluate themselves negatively. Socially anxious subjects have been found to negatively evaluate their social abilities (Alden et al., 1994). Lake and Arkin (1985) found that subjects who were higher in social anxiety rated positive feedback from evaluators as less accurate than participants who were low in social anxiety. Cacioppo, Glass, and Merluzzi (1979) investigated heterosocial (dating) anxiety, using the SAD as their measure of social anxiety and male subjects in anticipation of an interaction with a female as the anxious situation. Men who were high in social anxiety produced more negative self-statements and evaluated themselves more negatively. Clark and Arkowitz (1975) also found that subjects who were high in social anxiety rated themselves more unfavorably on a social encounter and that they had lower self-esteem.

Measures of self-esteem often have been used as an indicator of self-evaluation. Jones, Briggs, and Smith (1986) found negative correlations (ranging from -.52 to -.58) between self-esteem and various measures of social anxiety. Leary also found that social anxiety is negatively correlated with self-esteem ($r = -.36$, Leary & Kowalski, 1993; $r = -.18$, Leary, 1983). There may be gender differences. Endler,
Edwards, and Vitelli (1991) found no correlation between self-esteem and social evaluation anxiety for males, but a correlation of \(-0.42\) for females.

As previously stated and as is evident in this section, research in social anxiety comes under many different labels. Negative correlations also have been found between shyness and self-esteem \((r = -0.48, \text{Zimbardo, 1977;}}\ r = -0.51, \text{Cheek & Buss, 1981})\). Furthermore, in the communication apprehension literature, several researchers have found a moderately strong negative correlation (ranging from \(-0.48\) to \(-0.72\)) between self-esteem and communication apprehension (McCroskey, Daly, Richmond, & Falcione, 1977; and see review by McCroskey, 1977). Communication apprehension is "an individual's level of fear or anxiety associated with either real or anticipated communication with another person or persons" (McCroskey, 1977, p. 78); a construct that may be considered to be a subtype of social anxiety (Leary & Kowalski, 1995).

Thus the negative relationship between the two constructs seems to be a stable finding. Evaluating oneself unfavorably may lead to the expectation that others will evaluate you negatively as well (Leary & Kowalski, 1995). Thus, it may be the case that low self-esteem leads to an increased fear of negative evaluations from others, which leads to increased social anxiety. Future research should investigate if the fear of negative evaluations is a mediator between low self-esteem and social anxiety.

**D. SELF-REINFORCEMENT AND SOCIAL ANXIETY**

Self-reinforcement is the final element in self-regulation theory (Kanfer & Karoly, 1972). Upon appraising behavior, if it meets with the preset goal, self-reinforcement may or may not take place. A low frequency of positive self-reinforcement may be an antecedent of social anxiety. Rehm and Marston (1968) placed male college students who reported social anxiety into one of three therapy conditions. The condition that is relevant to this discussion involved increasing the client's rate of self-reinforcement. The greatest improvement was found for subjects in this condition. Individuals high in social anxiety may engage in a low frequency of self-reinforcement. More studies need to be conducted in this area to draw more definite conclusions about this relationship. Related research has shown that individuals who are high in social anxiety report less positive thoughts and more negative thoughts (Bruch, Mattia, Heimberg, & Holt, 1993).

**E. IMPLICATIONS FOR TREATMENT**

According to the research discussed in the foregoing text, individuals who are socially anxious do not expect to attain their desired goals, tend to be more aware of themselves as social objects, have lower self-esteem, and reinforce their behavior at a lower rate compared to normal controls.
Additionally, socially anxious individuals may engage in perseverative self-monitoring and self-evaluation. Cognitive behavior therapy may be a viable option to address their cognitions regarding the expectation of not attaining their goals. The negative self-evaluations based on research asking socially anxious individuals to rate themselves, as well as correlational studies that have shown a relationship between self-esteem (as an indicator of self-evaluation) and social anxiety can be addressed in cognitive behavior therapy as well. Finally, the Rehm and Marston (1968) study, discussed before, has provided evidence for the efficacy of increasing the self-reinforcing behavior of socially anxious individuals. These are all issues that may be addressed in therapy; however, there are many other variables related to the particular individual that may affect the process and outcome of the psychotherapy.

V. SELF-REGULATION AND DEPRESSION

Similar to the preceding section on social anxiety, various aspects of the self-regulation of behavior may contribute to depression for some individuals. One of the earliest and most well-known cognitive approaches to depression is the work of Beck (1967), which stressed dysfunctional cognitions. These dysfunctional cognitions form part of the self-regulation model for depression. At the goal setting stage of self-regulation, depressed people may set unrealistically high standards; at the self-evaluation stage, depressed people may be prone to cognitive distortions; and at the self-reinforcement stage, depressed people may have low rates of positive reinforcement or alternatively may negatively reinforce their behavior (Matthews, 1977; Rehm, 1977). These areas will be discussed separately.

A. GOAL SETTING AND DEPRESSION

Beck (1967) theorized that depressed individuals set unrealistically high standards for themselves: The individual would be unable to achieve his or her goals and depression would result. It has been found by several researchers that depressed individuals have higher expectations for themselves than nondepressed individuals (Golin & Terrell, 1977; LaPointe & Crandell, 1980; Nelson, 1977). However, in a laboratory study, Nelson and Craighead (1981) investigated the difference between set goals and actual performance on a task in depressed and nondepressed participants. No differences were found. Dysphoric students have been found to set lower goals than nondysphoric students (Ahrens, Zeiss, & Kanfer, 1988). Furthermore, Kanfer and Zeiss (1983) found that depressed participants set slightly lower standards than nondepressed participants. Additionally, they found that depressed participants felt that they had higher standards than
they were capable of achieving. Thus, perceived ability to attain goals may be the important factor.

Another relevant finding is that dysphoric individuals believe that others have high goals for them. Socially prescribed perfectionism is the tendency to believe (either correctly or incorrectly) that significant others have perfectionistic standards for the self. A positive association between socially prescribed perfectionism and depression has consistently been found (Hewitt & Flett, 1991, 1993; Martin, Flett, Hewitt, Krames, & Szanto, 1996). Similar to the socially anxious participants discussed previously, dysphoric participants also scored higher on the socially prescribed perfectionism measure than did controls. When asked to objectively rate the goals of others, however, their ratings also were no different than nondysphoric participants (Alden et al., 1994).

B. SELF-MONITORING AND DEPRESSION

Depressed individuals may engage in perseverative self-monitoring and self-appraisal which may contribute to and exacerbate their depression. As with socially anxious individuals, the relationship with self-consciousness has been investigated. The tendency to be aware of oneself as a social object is measured by the Public Self-Consciousness subscale of the Self-Consciousness Scale (Fenigstein et al., 1975). Kuiper, Olinger, and Swallow (1987) found that depressed participants reported higher levels of public self-consciousness than nondepressed undergraduate participants. Similarly, in a sample of undergraduate participants, Flett, Blankstein, and Boase (1987) found that public self-consciousness was a significant predictor of depression. Further support of the relationship between public self-consciousness and depression comes from a clinical study that investigated depression among a sample of pregnant women (Kitamura, Sugawara, Sugawara, & Toda, 1996). Women who were higher on a self-report measure of depression also were found to be higher on public self-consciousness.

The positive relationship between public self-consciousness and depression appears to be a stable finding. It is unclear whether depression causes an increase in self-consciousness or whether an increased level of self-consciousness predisposes one to depression. Some research has provided evidence to suggest that depression may lead to increased self-consciousness (Hull, Reilly, & Ennis, 1991), but further research needs to be done.

C. SELF-EVALUATION AND DEPRESSION

Beck (1967) stated that depressed individuals evaluate themselves negatively due to their unrealistic goal setting behavior. Although depressed people may not set goals that are higher than other individuals, they seem to set goals that they do not perceive themselves as capable of achieving.
Bandura (1977) defined self-efficacy as the perceived ability to behave in a manner consistent with achieving goals. Depressed individuals may have low self-efficacy. Depressed participants have been found to judge themselves as less self-efficacious (Kanfer & Zeiss, 1983). A large discrepancy has been found between goals set and efficacy ratings for dysphoric participants (Doerfler & Aron, 1995). However, no difference was found between the evaluation of performance between dysphoric and nondysphoric participants (Doerfler & Aron, 1995).

Another important element in the onset or maintenance of depression may be the frequency of self-appraisal. People may vary in the rate at which they evaluate their behavior. Dysphoric individuals have been found to engage in more frequent self-evaluation than nondysphoric individuals (Alden et al., 1994). Thus, a combination of negative self-evaluation and a high frequency of self-evaluation may contribute to depression. This research was conducted with undergraduate students. Further research should compare clinical populations with normal controls. Additionally, a within subjects design could be used to compare individuals while depressed and then later on when their depression has lifted.

D. SELF-REINFORCEMENT AND DEPRESSION

According to Beck (1967), depressed individuals engage in more self-criticism than normals due to their negative evaluations, which resulted from their unrealistic goal setting. This self-criticism can be classified as negative self-reinforcement. Lewinsohn (1974) postulated that depression was primarily due to a lack of positive self-reinforcement that resulted from a lack of social skills. It seems intuitive that depressed people would engage in less self-reinforcing behaviors; however, this was not confirmed empirically. Nelson and Craighead (1981) did not find a significant difference between depressed and nondepressed individuals on frequency of self-reinforcement.

Depressed individuals may benefit from therapy aimed at increasing positive self-reinforcement even though they may not be lower in the frequency of self-reinforcement than normal controls. Future research should determine the effect of therapy aimed at increasing positive self-reinforcement. Additionally, there may be a qualitative difference in the type of self-reinforcement.

E. IMPLICATIONS FOR TREATMENT

According to the research, depressed individuals do not believe that they will attain their goals. As with socially anxious individuals, cognitive behavior psychotherapy may address this issue. Also similar to socially anxious individuals, depressed individuals may engage in perseverative
self-monitoring and self-evaluation, which may be addressed in psychotherapy. Within the addictive behavior research, "obsessive–compulsive self-regulation" was the recommendation for the prevention of self-regulatory failure and the attainment of goals (Kirschenbaum, 1987). Research in social anxiety and in depression seems to provide support for the assertion that "obsessive–compulsive self-regulation" may be part of the problem in socially anxious and/or depressed individuals.

Rehm (1981) provided a treatment program for depression that targets each aspect of the self-regulation process. Clients are to set attainable goals, to record experiences in such a way as to minimize the possibility of cognitive distortion, and to evaluate their behavior in the therapy session so that self-blame is discouraged and credit is given for successful behavior.

Rehm (1981) evaluated the various aspects of the self-control therapy. A self-monitoring condition, self-monitoring plus self-evaluation condition, self-monitoring plus self-reinforcement condition, a full self-control treatment condition, and a waiting list control were compared. All treatment conditions were found to be superior to the waiting list control condition; only minor differences were found between the remaining groups. Roth (1982) compared the self-control treatment with antidepressant medication treatment and found that both methods of treatment reduced depressive symptomatology, but the combination of the two approaches produced the best results. Evaluation of the participants at a 3-month follow-up showed that treatment gains were maintained. The self-control treatment approach also was compared with cognitive therapy (Thomas, Petry, & Goldman, 1987). These forms of psychotherapy were found to be equally effective in alleviating depression. Generally, good results have been obtained with the self-control treatment model (Bandura, 1997).

VI. SELF-REGULATION AND THERAPY

One of the aims of any form of psychotherapy is to enhance a client's self-regulatory behavior. Rogers (1954) emphasized the self in psychotherapy and believed that changes within the self would occur with the awareness of feelings. Successful psychotherapy involved greater acceptance of the self (i.e., greater congruence between the actual self and the ideal self). Although Rogers did not use specific, directive intervention techniques aimed at changing the self, other forms of psychotherapy use direct training in self-reinforcement and self-control (Kanfer, 1970). These concepts were discussed in the addictive behaviors section. Tools specifically aimed at improving self-regulation were discussed in the health psychology section. Cognitive behavior therapy may be effective in address-
Tests of a self-regulatory form of therapy with different populations than discussed so far have shown that it is an effective form of treatment. A self-regulatory model of treatment was found to decrease encopretic (i.e., involuntary elimination) behaviors (Grimes, 1983). A combination of self-monitoring and self-reinforcement techniques was found to decrease hyperactive behaviors and increase academic performance in hyperactive boys (Varni & Henker, 1979). Because self-regulatory therapy has been effective in impulse control, it is thought to be well suited to the treatment of borderline personality disorder clients, although no empirical research has confirmed this assertion (Kanfer, 1970; Westen, 1991).

Skills attained in therapy are especially important when one considers that the client will eventually stop receiving therapy. It is important that a client can set attainable goals, monitor his or her behavior, and evaluate his or her behavior. If there is a discrepancy between the goal and the behavior, an attempt should be made to reduce it. Finally, the client should engage in the reinforcement of positive behaviors.

VII. SUMMARY AND CONCLUSIONS

The self-regulation of human behavior involves setting goals, and monitoring and evaluating behavior and thoughts. An attempt is made to reduce discrepancies between standards and behavior. If goals have been achieved, a person may engage in reinforcing thoughts and/or activities (self-reinforcement). Problems in the self-regulation of behavior or thoughts may result in distress. It may be that goals are set at levels that are too high or at levels that an individual does not believe he or she is capable of achieving. An individual may engage in perseverative self-monitoring and self-evaluation or, alternatively, may disengage from the monitoring of behavior. Furthermore, positive behaviors and cognitions may not be reinforced.

Several areas of clinical psychology (i.e., addictive behaviors, social anxiety, and depression) were discussed in terms of failure in self-regulation. In the addictive behaviors field, a main cause for failure in self-regulation is the disengagement of self-monitoring. An individual ceases monitoring and evaluating behavior, and engages in behaviors that deviate from goals. Therapy can involve raising a client's self-regulatory efficacy and/or having the client engage in constant and detailed self-monitoring.

Contrary to the addictive behavior research, socially anxious individuals and depressed individuals have been found to engage in perseverative self-monitoring and self-appraisal. Socially anxious and depressed persons also may set goals that they do not believe they are capable of achieving.
Furthermore, socially anxious individuals may evaluate themselves negatively and engage in less self-reinforcement. Socially anxious and/or depressed individuals may benefit from cognitive behavior therapy, where they are to set attainable goals, evaluate behavior in a therapy session so that it is done fairly, and self-reinforce positive behaviors.

Thus, principles of self-regulation theory can be used in therapy. Individuals who exhibit the Type A behavior pattern also can benefit from such self-regulatory therapy. A main component is to train the client to self-monitor behavior so as to identify states of increased arousal. Type A behavior pattern individuals may also benefit from the use of self-regulatory tools such as meditation, hypnosis, relaxation therapy, and biofeedback. These self-regulation strategies also have been shown to be effective in the management of various health problems (e.g., hypertension). Therapy based on self-regulation theory is appealing because it trains clients or patients to do something for themselves, and these skills can be used and continued after the therapy has ended.

Future research should further investigate the relevance of goal setting, self-monitoring, self-evaluation, and self-reinforcement in addictive behaviors, the management of health problems, depression, and social anxiety. Much of the research that has been done has been conducted with the use of an undergraduate population and should be replicated with clinical populations. Research in self-regulation in clinical psychology should investigate the relationship between each aspect of self-regulation (i.e., goal setting, self-monitoring, self-evaluation, and self-reinforcement) and the particular area under investigation (e.g., social anxiety). The efficacy of therapy aimed at these areas should be investigated and designed such that it is possible to determine the effects of targeting only specific aspects of self-regulation in therapy (e.g., only self-monitoring) as well as the effects of targeting all aspects of self-regulation in therapy.

ACKNOWLEDGMENT

Preparation of this article was supported, in part, by Social Sciences and Humanities Research Council Grant No. 410-94-1473 to Norman S. Endler.

REFERENCES


596

PART III. INTERVENTIONS AND APPLICATIONS OF SELF-REGULATION


