NAEP TESTING FOR

TWELFTH GRADERS: MOTIVATIONAL ISSUES

Jere Brophy and Carole Ames

Michigan State University

A paper prepared for the National Assessment Governing Board

September, 2005

Executive Summary

This commissioned paper draws on theory and research on motivation in education to identify principles that might guide the National Assessment Governing Board's decision-making concerning twelfth-grade NAEP. It draws on three major areas of motivational theory to analyze the motivational challenges that need to be faced and identify principles that might guide efforts to address these challenges. It summarizes the motivational aspects of both the negative case (abandon twelfth-grade NAEP testing) and the positive case (continue or expand it). In elaborating the positive case, it identifies several strategies that might be employed in recruiting twelfth graders not only to participate in NAEP but to engage in it with optimal motivational orientations. Most of these strategies would require departures from traditional NAEP procedures, however.

Given the nature and purposes of the NAEP program, successful implementation of twelfth-grade NAEP would require not only recruiting an appropriate sample of students to participate, but enticing these students to do so with high levels of engagement that feature desired motivational orientations. In particular, it would be important to encourage the students to perceive participating in NAEP as a worthwhile thing to do, to feel that they are acting autonomously when choosing to do so, and to understand the purpose.

This is a tall order, given the content of NAEP, the traditional NAEP procedures, and the attitudes of most twelfth graders toward tests and test taking. Consequently, the most direct extrapolations from our motivational analysis favor the negative case (drop twelfth-grade NAEP testing). Neither students nor school personnel are likely to perceive value in NAEP participation. Tests are not intrinsically interesting to most

students, and although some may value rewards that successful test performance might bring, they usually do not value the process of test taking itself. This is especially true of high school seniors in their spring term, who are disengaging from the series of evaluation hurdles that is built into our high school culture. Because NAEP does not align with their school's curriculum and does not lead to feedback that would allow students to improve their school performance levels, and because there are no prospects for rewards either for participating in the assessment or for attaining some qualifying score, there is no rational incentive for students to participate. From their perspective, there is little or no potential benefit, but there are potential costs, especially for students for whom NAEP assessment presents the prospect of yet another set of experiences with frustration and failure. Also, because the NAEP program is low in national visibility and does not yield outcomes that have direct applicability or consequences for local schools or districts, there is no reason for school personnel or the students' families to value their school or district's participation in the NAEP program. These considerations suggest that the NAEP assessment of twelfth graders faces daunting motivational obstacles that would be difficult to overcome, so efforts to do so are not likely to be successful.

However, the same motivational principles that portend negative outcomes if traditional NAEP procedures are used to provide a basis for predicting more positive outcomes if traditional procedures could be changed to incorporate incentives to motivate participation and test preparation and administration procedures designed to elicit high levels of engagement. By shifting from a strategy of randomly selecting students in a school to a strategy of filling slots in a much smaller but highly representative sample, the

NAEP program might be able to afford most if not all of the suggested strategies and at the same time end up with data that are better suited to extrapolating national trends.

Most of the suggested strategies would be designed to encourage students to see value in participation in NAEP, and thus to accept invitations to do so. One strategy is to offer incentives, which might include paying students for participating or offering training in test-taking skills that would serve them well in the future. Other strategies would involve appealing to students' social and civic identities, by depicting participation in the NAEP program as an opportunity to help the test developers shape future tests and provide the government with important information about national trends (drawing a parallel to the families who participate in Nielsen television viewing surveys); appealing to students' identification with peers, school, or the community (by depicting the invitation to participate in the NAEP program as an honor for the school and community); and by depicting participation as an opportunity for the students to represent their school, their community, and even their extended peer cohort (i.e., America's teenagers) in ways that will bring them credit.

Other strategies might focus on enhancing the interest value of participating in the NAEP (emphasizing that the NAEP assessments are more interesting than most tests and that participation will include training in test-taking skills). Still other strategies would focus on reducing the perceived cost of participation to students, by reducing the actual costs in time and effort (if necessary) and by minimizing potential psychological costs (fears of failure or embarrassment).

In addition to these value-enhancing strategies, emphasizing that participation is voluntary and providing training that emphasizes self-regulation of one's test-taking

strategies would encourage students to engage in the assessment with a sense of autonomy. Encouraging students to do their best and arming them with strategies for doing so, while at the same time avoiding any suggestion that the tests measure ability or that students' scores will be made public and compared with those of peers, would encourage them to approach with a sense of competence. Finally, adjustments could be made in the conditions of testing, the instructions to the students, and the role of the proctor that would make students more comfortable in the testing situation and more likely to respond to all of the items for which they possess relevant knowledge Together, these strategies would encourage both the high rates of participation and the high levels of engagement that are needed to support the validity of NAEP results as bases for drawing inferences about national trends in student achievement.

In summary, our analysis indicates that NAEP assessments, as conducted in the past, offers nothing of objective value to participating students, their schools, or their communities, so that twelfth-grade NAEP probably should be dropped if NAEP policies require persisting with these same procedures. However, to the extent that these policies would allow for incorporation of the strategies suggested in the section of the report that presents the positive case, there may be reason for continuing or expanding twelfth-grade NAEP, or at least, conducting pilot efforts to assess the effectiveness of these suggestions.

This paper, commissioned by the National Assessment Governing Board, draws on theory and research on motivation in education to identify principles that might guide decision making concerning NAEP assessments for twelfth graders. It draws on three areas of motivation theory, identifies motivational issues that make twelfth-grade NAEP assessments problematic, and suggests relevant motivation principles that have implications for decision making. The implications section takes two approaches, first making the case for not testing twelfth graders, and second suggesting strategies that might be employed to help address the challenges involved in twelfth-grade NAEP.

Motivational Analysis of the Problem

Tests are designed to sample what the test takers know and can do in the domains identified by the tests. The underlying logic of testing, and especially testing programs such as NAEP that are designed to develop normative information about groups, usually includes the assumption that individuals' test scores are valid estimates of their true levels of knowledge or skill in the domain (with the exception of minor inaccuracies due to random errors). Implicit in this logic is the assumption that test takers are optimally engaged in the task: intending to do the best they can, reading the questions carefully and responding thoughtfully, and so on. If they are not optimally engaged, or if they take the test under seriously suboptimal conditions, their test scores will represent underestimates of their true scores.

In addition to physical and psychological problems (fatigue, drugs, clinical depression, etc.), suboptimal engagement in test taking can result from a variety of

motivational problems (e.g., anxiety and fear of failure continually distract students from test-relevant information processing and decision making; low frustration tolerance or catastrophic reactions to failure lead to cessation of responding or random, rather than thoughtful, responses to test items). Thus, one challenge inherent in testing situations is arranging testing conditions to encourage optimal task engagement during test taking, so that the obtained scores represent unbiased estimates of true levels of knowledge and skill.

In addition to this engagement challenge, the twelfth-grade NAEP program faces a participation challenge. Although much testing is required as a part of schooling or a condition of employment, participation in the twelfth-grade NAEP program is voluntary. Consequently, this program faces the challenge of recruiting students to take the tests in the first place, not just attempting to optimize their engagement during test taking.

Research on motivation in education is directly relevant to these challenges because it is all about choosing whether or not to participate in an activity and the factors that influence the quality of such participation. Motivation is reflected in people's choices of activities, the amount of effort they invest in these activities, the kind of cognitive and self-regulatory strategies they apply, and their persistence over time and perseverance in the face of difficulty or frustration. In addition, motivation is about why and how students engage in the activities they choose. The reasons for their choices and the goals they pursue as they engage in the activity affect the quality of their engagement and ultimately the outcomes of their effort investments. So, assessment goals and procedures that include motivation principles should aim not merely to elicit compliance

or agreement to participate, but also to induce optimal test-taking attitudes, strategies, and processes.

Three Relevant Motivational Theories

Research on motivation in education has produced a range of useful theories and models, although many of these are more applicable to learning situations than to testing situations. For purposes of this paper, we will draw primarily on three programmatic lines of theory and research that have direct relevance to the participation and engagement challenges faced by the twelfth-grade NAEP program. These are expectancy x value theory, self-determination theory, and goal theory.

Expectancy x Value Theory

A great deal of what researchers have learned about motivation is related to student expectation of success and the perceived value of a task or activity (Brophy, 2004; Feather, 1982; Pekrun, 1993; Wigfield & Eccles, 2000). Choosing to engage in an activity and expend effort on it depends on whether an individual expects to be able to perform the activity successfully (and thus receive rewards for a successful performance) and whether the individual values those rewards, sees the activity as worthwhile or personally relevant, and/or values the processes involved in performing the activity itself.

Effort investment is the product of expectancy and value factors. Thus, no effort will be invested in an activity if one of these factors is missing entirely. People do not choose to invest effort in activities that they do not value or do not see as worthwhile, even if they know that they can perform these activities successfully. Nor do they

willingly invest effort in even highly valued activities if they believe that they have no chance for success.

With respect to the recruitment of twelfth graders to participate in NAEP, the value factor is fundamental. Successfully addressing the value aspects of motivation involves providing compelling answers to questions such as, Do I care about this activity? Is it worthwhile? What are the benefits from engaging in it? Are these benefits sufficient to justify the time, effort, or other costs involved?

Strategies must be developed to induce twelfth graders to value participation sufficiently to motivate them to agree to participate. They must view, or be induced to view, taking the test as a worthwhile thing to do. Value is related to four dimensions (Eccles & Wigfield, 1985):

1. Attainment value which is the importance of attaining success on a task in order to affirm one's self-concept or fulfill needs for achievement or prestige
2. Intrinsic or interest value is the enjoyment one gets from engaging in the activity
3. Utility value is the role the activity may have for advancing one's progress toward larger or longer term goals
4. Cost refers to the time, effort, or other resources required to participate in the activity, as well as foregone opportunities to use those resources for other purposes.

Eccles and Wigfield's model suggests several approaches for responding to students' value questions regarding NAEP participation. In addition, Brophy (2004) has addressed the value aspects of motivation at considerable length in a text on motivation in education.

The expectancy aspects of motivation are rooted in students' beliefs about their prospects for meeting the challenges embedded in an activity and thus completing it successfully. Bandura (1997, p. 3) refers to these beliefs as self-efficacy perceptions, which he defines as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments." Individuals with positive self-efficacy perceptions believe that they can accomplish what the situation calls for, but people with low self-efficacy perceptions are less certain. When self-efficacy perceptions are high, people are likely to approach achievement situations with confidence and engage in them willingly and persistently. However, if they doubt their capabilities for succeeding, they are likely to try to avoid the situation, or if that is not possible, to give up easily when they encounter frustration or failure.

The expectancy aspects of motivation are less salient in twelfth-grade NAEP compared to other testing conducted in school contexts, because students' scores will not become public and will not affect their school grades or future opportunities for employment or higher education. Even so, other attitudes and beliefs, relating to the students' prior achievement histories and perceptions of their own test-taking capabilities, are likely to influence their thinking about whether or not to participate in the assessment and, if they do agree to participate, the quality of their engagement. Consequently, the National Assessment program should take steps to minimize the degree to which low self-efficacy perceptions, failure expectations, and other expectancy-related problems cause some students either to refuse participation or to participate but perform below their capabilities.

Expectancy-related beliefs also affect students' thinking about the costs vs. benefits of putting forth one's effort. Perceived costs are related not only to the amount of time and effort required, but also to anticipated consequences (e.g., social) of participating or spending too much time on the task. Students who see value in an activity in terms of personal relevance, social norms, or future goals, and who see the costs of participation as not excessive, are more likely to choose to participate and engage actively in the full scope of the activity.

Self-Determination Theory

The self-determination theory of Deci and Ryan (1985, 2002) begins with the observation that when people are motivated, they undertake specific actions to accomplish a goal. Their motivated action may be either self-determined or controlled. To the extent that it is self-determined, it is experienced as freely chosen and emanating from oneself, not under pressure from some external force.

Self-determination theory emphasizes the importance of satisfying one's needs for autonomy (deciding for oneself what to do and how to do it), competence (developing and exercising skills for success), and relatedness (affiliating with others). In other words, people are inherently motivated to feel connected to others within a social milieu, to function competently in that milieu, and to feel a sense of personal control while doing so.

Self-determination theory emphasizes the importance of subjective perceptions of control. The perception of self-control and autonomy is critical, even if extrinsic incentives are in effect or if our behavior is constrained in various ways. For example,

even though most K-12 students are required to attend school, instruction that allows for socializing and autonomy in learning will benefit student participation and engagement. Deci and Ryan (1994) identified three factors that promote self-determination in classrooms and also apply to testing situations:

1. Provide students with meaningful rationales that will enable them to understand the purpose and personal importance of the test;
2. Acknowledge students' feelings when it is necessary to require them to do something that they do not want to do (by letting them know that you are aware of their feelings and taking time to explain why the requirement is needed); and
3. Manage the situation using a style that emphasizes choice rather than control.

Goal Theory

Goal theory focuses on qualitative aspects of students' choices and engagement, particularly their beliefs about the reasons for engaging in achievement-related behavior (Ames, 1992; Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002; Kaplan and Middleton, 2002; Maehr & Midgley, 1996). These reasons or purposes have been contrasted as directed toward mastery or performance goals. Students who engage in achievement activities with mastery goals focus on acquiring the knowledge or skills that the activities are designed to develop. In contrast, students who engage in the activities with performance goals view achievement activities as tests of their ability to perform, as opportunities to demonstrate their ability, or as challenges to their ability (rather than as opportunities to learn). Performance-focused students seek to demonstrate high ability

relative to others or avoid being viewed as low in ability. Their emphasis is on maintaining an image rather than on learning.

Positive learning and motivational outcomes are most likely when students are operating with mastery goals. In short, outcomes are optimized when students are focused on mastering the content and skills involved in learning activities rather than thinking about competing with peers or worrying about how their performance will be perceived by others.

When ability comparisons are frequently made or easily inferred, students may become more concerned with preserving their sense of self-worth than with mastering the curriculum (Covington, 1992). This can lead to face-saving, but ultimately counterproductive, reactions such as pretending to understand when they do not, refusing to ask for help, or engaging in self-handicapping strategies (such as not studying for a test or procrastinating until the last moment) that position the students so that if they do poorly, they can blame their failure on something other than their own lack of ability.

In general, goal theory has more relevance for learning than testing situations, especially if the test, like the NAEP, does not yield feedback that directly relates to what students have been learning or provide information about remediation. Even so, if students view the test as a threat to their own sense of competence or self-worth, they may adopt a performance goal. And, students who have doubts about their ability or anxiety about their performance may engage in strategies to reduce the importance of the test and implications of failure. For example, they may decline to participate or may try to convince others to not participate, not try to do their best, or undermine the purpose of the test. Although high achieving students may welcome the opportunity to test their

knowledge and skills, low achieving students or those who doubt their ability to perform to their own standards will be more motivated to avoid participation and co-opt others to do the same.

It is important to make sure that testing conditions do not foster performance orientations. What is conveyed initially when recruiting students for participation and later when giving instructions in the testing situation should be free of any suggestion that the tests measure ability, that student performance will be compared or labeled, that individual results will be made public, and so on. All of the students' cognitive resources should be focused on meeting the demands of the test, so that they are not distracted by performance concerns, social comparisons, or self-doubts.

Other Motivational Considerations

All three theories converge on findings that pessimistic expectations, low self-efficacy perceptions, and performance goals need to be kept out of the situation or at least minimized as much as possible. These motivational conditions are associated with superficial information-processing strategies; loss of concentration due to performance anxiety, frustration, catastrophic reactions to failure, or other emotional problems; and undesired reactions such as just giving up, beginning to respond randomly so that one can pretend to have finished early, or deliberately procrastinating, breaking pencils, or otherwise handicapping oneself so as to be able to achieve "failure with honor." At the same time, students need to see the test as worthwhile, as having value, and as having benefits that outweigh the costs of time and effort. Other considerations that lie outside of these theories are also worth mentioning.

Developmental Trends

Liking and valuing of schoolwork and everything connected with it (including tests) tends to decline over the school years. At the same time, students become increasingly more selective of activities based on personal interests, perceived value, and perceived competence. Compared to younger students, high school students are more aware of social comparisons and thus more prone to developing self-worth concerns and performance goal orientations. Yet, high school contexts are more performance-oriented, with increased competition and social comparison (often including tracking) and with more frequent, normative, public, and higher-stakes evaluation. As a result, many students become focused on meeting minimal requirements and engaging in cost-benefit analyses with regard to choices of activities and allocation of effort to assignments and studying for tests.

Social Influences

Parents and the peer culture influence students' academic choices, behaviors, and goals. Parents are not involved in the day-to-day schooling of high school students but they do convey beliefs, values, and expectations that influence their children. For example, some parents place more value on performance—demonstrations of ability and achievement, public recognitions and awards, grades, and winning. Other parents place more value on the process of learning, taking on challenges, gaining competence, and improvement. These beliefs can influence students' choices and the quality of their engagement, attitudes, and self-worth.

The peer culture is at least as important in influencing students' motivation. Peers' attitudes and beliefs often shape those of peer group members, for good or ill. In some schools or classes, a dominant ethos of identification and cooperation with authority figures may support efforts to recruit participation in NAEP testing. In others, however, a contrasting ethos may necessitate preventive strategies such as making sure that commitments to participate are conveyed privately rather than publicly or emphasizing financial incentives (rather than providing service) when seeking to enlist participants.

The Negative Case: Drop Twelfth Grade from the NAEP Program

Tests are not intrinsically interesting to most students and although some students may value the reward that successful test performance might bring, they usually do not value the process of test taking itself. Given that the NAEP does not align with their school's curriculum and that there are no prospects for rewards either for participating in the testing or for attaining some performance level, there is no rational incentive for students to participate. From their perspective, there is little or no potential benefit, but there are potential costs. For all students, there is the time needed to participate in the assessment, which might have been allocated to some other purpose of more importance to them. Furthermore, for many students, especially those with low expectation problems, NAEP presents the prospect of yet another set of experiences with frustration and failure.

Lack of enthusiasm for NAEP participation is especially likely among high school seniors. The culture of the senior year has become an established "rite of passage"

through a system and culture that has overemphasized evaluation. By the spring semester, most of the evaluation hurdles have been completed and the high school culture moves to minimize requirements and instead focus on conclusions. The performance-oriented context of many high schools creates motivation to maximize outcomes with minimal effort, make choices based on the costs involved to achieve personal gains, and resist external control and imposed requirements.

In addition, there is little or no incentive for schools to invest resources and effort into encouraging, recruiting, or preparing students to participate. NAEP is a national assessment program that lacks direct applicability or consequences for local districts or schools; it is not a research program that yields findings that could provide input to curricular or instructional decisions. NAEP also has low national visibility, and thus little perceived value or importance to parents or the community. Consequently, there is little basis for claiming that their children's participation might benefit education in the future.

In summary, students have little reason to choose to take the NAEP tests, to fully engage in doing so, or to use strategies likely to optimize their performance. From their perspective, there is no apparent value in taking the test. It has no importance to students, their families, or even the broader school or district community. Consequently, the test lacks apparent utility or instrumental value, yet entails costs (minor for many students but potentially major for many others).

Taken together, these considerations point toward the conclusion that NAEP assessments of twelfth graders faces daunting motivational obstacles that are difficult to overcome, so that efforts to do so are not likely to be successful (i.e., to yield the desired rates of participation and quality of engagement).

The Positive Case: Potential Success Strategies

Despite the daunting challenges just summarized, the motivational literature does suggest principles and strategies that might be employed in efforts to recruit twelfth graders to participate in NAEP and maintain high-quality engagement when taking the assessments. Thus, it is possible to use this literature to make a case for continuing to assess twelfth graders, or at least, piloting some or all of the suggestions below and then deciding what to do.

In deriving principles and strategies, we focused on issues involved in motivating students to participate in the assessment program and to engage in it in ways that would enable them to perform up to their capabilities and thus yield maximally valid scores. We did not attempt to take into account other issues such as whether a strategy would be considered appropriate given the purposes and goals of the NAEP program, whether sufficient funds would be available to pay the cost of implementing the strategy, or whether NAEP already employs or would be able to hire people capable of carrying out the strategy effectively. Judgments on these issues are best made by NAGB members.

However, we do have one major suggestion that we believe might address assessment design, cost, and motivation issues simultaneously: Rather than trying to recruit a random sample of twelfth-grade students within each school, confine the assessment to a carefully selected subset of students stratified by prior achievement levels (and perhaps also by sex or race/ethnicity if that were considered desirable). For example, once a school has agreed to participate in the NAEP, the cumulative GPAs of its twelfth graders (either across all subjects or specifically in the subject to be tested)

would be analyzed to identify its distribution characteristics, and a plan would be developed for drawing a sample of twelfth graders that reflected those characteristics. As a simple example, in a school that had 200 seniors, 20% might have GPAs between 3.5 and 4.0, 20% between 3.0 and 3.5, 20% between 2.5 and 3.0, 20% between 2.0 and 2.5, 10% between 1.5 and 2.0, and 10% below 1.5. The sampling plan for such a school might call for testing 40 students, drawing eight students from each of the four highest GPA categories and four students from each of the two lowest categories. Within categories, random selection procedures would be used to determine which students in each GPA range would be initially identified as the intended sample, which would be identified as the first alternate to turn to in case of a refusal, which would be the second alternate, and so on. Selected students would be recruited individually, with the recruiters employing some of the motivational principles and strategies outlined below. Essentially, we recommend that NAEP use the sampling-with-replacement strategy that polling organizations use to ensure that their sample characteristics reflect the distributions that exist in the population at large.

Given the relatively low rates of participation by public school students in previous rounds of twelfth-grade NAEP assessments, it seems to us that a smaller but carefully recruited stratified sample would be more representative of the full population of the nation's high school seniors (and thus preferable from a normative assessment perspective) than the current samples that are biased toward the more compliant and higher achieving students. Testing a smaller sample also might free up at least some of the funds needed to pay for the strategies outlined below.

Recommended Principles and Strategies

The three lines of motivational theory and research that we have emphasized suggest the following basic principles for motivating students to participate in NAEP at a high level of engagement: encourage students to value such participation (see it as worthwhile) and take steps to minimize its potential costs to the students; emphasize individual choice and autonomy; and depict and administer the assessments in ways that avoid a performance goal orientation (e.g., avoid depicting the tests as measuring ability, as well as any comments that suggest competition or social comparison). Several potential strategies for implementing these principles are suggested below.

Create Utility Value by Offering Incentives

Since the NAEP is not directly connected to students' school curricula and the absence of familiar norms and individualized feedback make it impossible for students to set specific goals for their test performances, NAEP does not offer opportunities for students to experience attainment value. However, we can see ways to inject utility value and perhaps interest value into the experience.

1. Offer incentives. When activities are not meaningful, interesting, or personally relevant to students, external incentives may provide them with a reason for engagement. Monetary payment is the ideal choice of incentive because it is easy to administer and because money is attractive to all students.

If students are paid for their participation, it will be important to explain and implement the payment in ways that support high-quality engagement in test taking. One difficulty with extrinsic rewards is that they may be perceived by students as attempts to

control their behavior, which can lead to apathetic or even covertly resistant forms of participation. However, this difficulty can be minimized by designing the payment procedure with the following principles in mind.

First, the students should be invited but not pressured to participate. They might be told:

> This is an opportunity to participate in our nation's national assessment program, to provide feedback to the test developers, and to acquire some testing knowledge and experience that should be useful to you personally. It is not for everyone, however. It carries obligations for putting forth your best efforts when taking the test and providing thoughtful feedback later. If you choose to participate, keep this responsibility in mind. You are being paid for your time and best effort. If you choose not to participate, no problem.

Second, encourage students to perceive the money they receive as "pay for work" rather than "working for rewards." The payment is not contingent on test performance, but it does assume high quality engagement. This should appeal to students' sense of fairness and appropriateness in that they are being paid for doing a good job. This provides a reason for engagement but also allows them to maintain a sense of personal control and choice about participation.

Finally, make sure that the pay is sufficient to justify these perceptions. The amount should be perceived as sufficient compensation for the time and effort required, not a mere token (e.g., perhaps $10.00 per hour).

2. Offer training in test-taking skills. Another potential incentive that could be built into the participation experience would be one or more sessions in which students

are given training and tips in test-taking strategies. The training should be extensive enough to be of value to students not only in the short term when they take the NAEP tests, but in the future when they take tests that have implications for employment opportunities or college admittance. It should include attention to both cognitive and emotional self-regulation and test-taking strategies, and thus both minimize test anxiety and increase students' confidence about their ability to take tests proficiently.

Many schools already provide some training in test taking, but few also provide opportunities for debriefing following test taking. Participants might be given time to discuss the value of various strategies, share their experiences of implementing the ones they were taught and perhaps others they developed on their own, and so on. This would help students to internalize the strategies, and in the case of those who are prone to performance-avoidance goals, would encourage them to take more of a mastery goal orientation toward test taking in the future.

In summary, to enhance students' perceptions of the utility value of participating in the NAEP program, recruiters might suggest something like the following:

> Participation likely will have indirect benefits that will help you in the future when taking college entrance exams or employers' tests. It will include opportunities to learn test-taking skills and strategies before the tests are administered and to discuss their value afterwards. Also, NAEP usually doesn't give students feedback, but in this case you will have the opportunity to review and make comments on the test items with representatives of the NAEP program, and this will include opportunities to learn the best answer to each item.

Appeal to Students' Social and Civic Identities

In addition to or instead of building utility value into NAEP participation as described above, recruiters might encourage students to perceive participation in the assessment as carrying symbolic or identification value. These strategies would appeal to their identities as American high school students and as citizens providing service to the nation.

1. Opportunity to help the test developers and shape future tests. To appeal simultaneously to students' affiliation/belongingness needs, social responsibility needs, and power needs, recruiters might emphasize that participation in the NAEP and subsequent feedback sessions provides important service to the test developers and ultimately the nation as a whole. Students who participate in NAEP are like families who participate in the Nielsen ratings for television programs. This sample is the basis for drawing generalizations about the nation as a whole, in this case, about how students are doing in various subjects. Because the sample needs to be representative of the nation as a whole, participation by all kinds of students is needed, not just high achievers or other special groups.

Also, by providing thoughtful feedback (by writing it on response sheets and/or participating in focus groups), students can supply the test developers with much needed information about the test as a whole (quality of instructions, time allotted for each subpart) and about the clarity and appropriateness of individual items. This feedback will help the developers make the tests as good as they can be, thus maximizing their validity as measures of student learning in the subject areas.

2. Appeal to students' identification with peers, school, and community. Students sometimes will engage in a task that is not of great personal interest to them if they are part of a peer, school, or community culture that places importance on the activity. Motivation stemming from identifying with and valuing being a part of this extended culture may make up for an absence of personal interest or individually perceived value. Both NAEP recruiters and school administrators and teachers could encourage such motivation by portraying the school's selection for NAEP participation as an honor and opportunity ("We are excited to hear that our school has been invited to participate in the National Assessment of Educational Progress, our nation's . . . ").

To enhance the perceived value (in the eyes of both students and their parents) of both the NAEP program and the potential for participants to provide feedback that could make a difference, NAEP might consider establishing follow-up opportunities. For example, selected twelfth graders might be invited to a national conference on how to improve high schools, how NAEP assessments might be better linked to high school curricula, how to improve test-taking strategies, etc. Participation in the NAEP would be linked to the opportunity to be invited to this conference. The selection of representatives would be based on national criteria and not restricted to high test scorers or students at certain schools.

3. Opportunity to show what we know. School administrators and teachers could also portray participation in NAEP as an opportunity to "show them what we know"—to represent their school, their community, and even their extended peer cohort (i.e., America's teenagers) in ways that will bring them credit. It could be noted that, these days, critics are always depicting schools as not doing a good job and students as not as

well informed as previous generations. Participation in NAEP then would become an opportunity to refute these critics, and students might be mobilized to do their best and show what their generation can do.

In implementing this strategy, it would be important to speak in terms of showing what we know or what we have learned rather than showing how smart we are or proving our abilities. The former language is consistent with mastery goal orientations, but the latter language suggests performance-goal orientations and subtly encourages students to think of the NAEP as tests of ability rather than acquired knowledge.

Enhance the Interest Value of Participation in NAEP

Even though test taking is not an inherently interesting activity for most students, some of the symbolic/identification strategies just described would add perceived interest value to NAEP participation. This might be enhanced through provision of information about the tests themselves:

> NAEP tests are more interesting than most tests because besides 'mark the most correct choice' items, there are constructed response items that ask you to generate an answer and explain your thinking in a short paragraph. Also, even though the tests contain some items that few if any of you are likely to get correct, they do not just begin with simpler items and then get increasingly harder. Sometimes, a question about something that you don't know anything about at all will be followed by a question about something that you do know or can figure out on the spot. Finally, instead of just asking you questions about a single instructional unit or group of lessons, the test items cut across all areas of the

curriculum. So, you will find that NAEP questions are challenging but unusually interesting, and you are likely to find that you have a lot of feedback to give to the test developers. The training in test-taking strategies that you will receive before you take the tests will include attention to these features of the NAEP tests and the best ways for you to respond to them.

Reduce the Perceived Cost of Participation to Students

In deciding about whether or not they want to participate in the NAEP program, students will consider not only the perceived value but also the perceived costs of doing so. Consequently, recruiters should encourage students to view these costs as minimal.

1. If necessary, reduce the actual costs in time and effort. If participation is likely to be perceived as involving lengthy testing, fatigue, etc., NAGB might consider taking steps to reduce the time demand on any individual student. Such steps might include reducing the length of the overall test or using different samples of students to respond to different subtests.

2. Minimize fears of psychological costs. For students in general and especially for students with chronic patterns of low achievement or test anxiety, it will be important to portray the NAEP testing experience as an opportunity and not a threat. At minimum, this will mean emphasizing clearly that NAEP assesses not ability but knowledge and skills acquired at school; that the range of item familiarity/difficulty is such that most if not all students are likely to encounter at least some items on which the best they can do is make an educated guess (and this is fine, because the test is designed to probe the limits of knowledge, not just

assess basic understandings expected of everyone); and the training offered prior to the testing will familiarize students with the kinds of items to expect and suggest strategies and tips for responding to them effectively.

Where appropriate, test preparation might also include principles and strategies that have been developed for minimizing test anxiety problems (Brophy, 2004; Hembree, 1988; Hill & Wigfield, 1984; Neveh-Benjamin, 1991; Wigfield & Eccles, 1989; Zeidner, 1998). Also, with students who are members of groups who might be expected to perform poorly on tests in general or tests in particular domains, test preparation might include components intended to counteract stereotype threat and minimize its negative effects on performance (Aronson & Steele, 2005;Cohen, Steele, & Ross, 1999; Croizet et al., 2001; Johns, Schmader, & Martens, 2005).

In addition to preparing students for the assessment through advanced training, the NAEP program should include instructions to test administrators that would minimize anxiety in the testing situation and inoculate students against catastrophic responses to failure. For example, they might explain that NAEP is one of those tests on which questions range from very easy ones that most students will answer correctly to very difficult ones that few if any students will answer correctly. So, students should not be surprised if they see questions about things they have never studied. They should just answer as best they can and move on, because the next question probably will be easier.

Foster Perceptions of Self-Determination

Some strategies already mentioned in other contexts are worth repeating here in the context of fostering perceptions of choice and autonomy. Much of this will be accomplished by making clear to students that their participation is voluntary. Even though participation might bring pay or other rewards and agreement to participate implies commitment to putting forth their best efforts, the choice is theirs. Thus, in this case, not only is it possible for students to feel self-determined in the testing situation, in point of fact their participation will be self-determined.

Additional support for feelings of self-determination might be provided by acknowledging students' feelings (e.g., "We understand that you are twelfth graders and tired of taking tests by now, but this is important to the nation and we will do all we can to make sure that it has value for you as well. But again, the choice is yours.").

Encourage Mastery Rather Than Performance Orientations

Much of this also would be covered indirectly in previously suggested strategies, but they are worth repeating here in the context of avoiding performance orientations. Anything that recruiters, test administers, or school personnel say about the NAEP should avoid any suggestion that these are tests of ability or intelligence, as well as any suggestion that individual students' scores will be made public or compared with those of peers. Instead, comments about NAEP should emphasize something like the following:

> The NAEP will be administered to a representative sample of our nation's high school seniors. The purpose is to generate information about what knowledge and skills have and have not been learned in the different subject areas. Your role is to put forth your best efforts and answer as many questions as you can. We will

give you tips about how to manage your time to allow you to get to as many questions as possible and bring to bear what you know about each subject.

Improve Testing Conditions

It would be possible to augment the current system for monitoring students' task completion and engagement. For example, each test segment might contain one or more questions designed to see if students are carefully reading the questions and responding thoughtfully to them. For these questions, the correct response choice would be obvious, so at least 95% of the students ought to get it correct. Data from students who failed to answer all (or perhaps all but one) of these questions correctly could be discarded and substitutes with similar GPAs could be recruited to take their places in the sample. Alternatively, these inserted questions might be used to determine at what point particular students became disengaged with the test and ceased responding or started responding randomly. Such data also could be used as part of a study designed to determine whether paying students for their participation (following guidelines suggested here) proves to be cost effective in generating higher participation rates or higher quality engagement than relying on non-monetary incentives or no incentives at all.

We do not think that such measures are necessary, however. In the first place, data from the 2000 and 2002 NAEPs administered to twelfth graders indicate that multiple choice items were attempted by 96% of students and constructed response items by 90%, with most failures to respond being classified as "not reached" or "omitted" rather than "off task." These response percentages might be as high as one could expect, given that the NAEP tests include items on content that some students have never had an

opportunity to learn (and items that are so difficult for some students that they cannot begin to form a response). In these situations, leaving the item blank is the "correct" response from the perspective of accurate assessment: The students have no relevant knowledge, so they put nothing down. It is not reasonable to expect all students to answer all items unless the test is confined to relatively basic levels of knowledge about content known to have been taught to all students tested.

The frequency of items classified as "not reached" is a concern, however. NAEP tests are supposed to be power tests, but we infer that they sometimes become de facto speed tests because sufficient time is not allowed for the group as a whole, or at least for slower students, to complete all items. To the extent that this is the case, the solution to the problem of "not reached" items lies in eliminating time pressures from the testing situation, not improving students' motivation.

We recommend addressing this problem as part of a larger effort to improve the testing situation in ways that reflect the motivational principles and strategies outlined above. We recommend testing the students individually or in small groups, in an appropriate setting outside of their classroom. This will help underscore the idea that participation in the NAEP program is a special event rather than just another school test, and it will make it easier for proctors to monitor individual students' engagement in the assessment. The proctor would explain that there is no time limit, so students can take as much time as they need to record and check their responses before turning in their booklets. When they finish, they should bring their booklet to the proctor for inspection (to make sure that they did finish, in the sense of responding to all items that they were able to respond to).

To set the stage for optimal engagement, proctors should conduct themselves as supportive resource persons, not authority figures there only to ensure compliance and prevent cheating. In the ideal case, the proctors would be the same people who provided the students with information about the NAEP tests and training in test-taking strategies prior to the testing session. At minimum, proctors might precede the testing with last-minute tips and review of test-taking strategies. This would include leading the group through some practice items and sample scoring rubrics for constructed responses (emphasizing that even if students do not know enough to answer some of these questions completely, they might earn partial credit by communicating what they do know), reminding them that a difficult or unfamiliar item might be followed by an easy or familiar one, and encouraging them to take as much time as they need.

Creating these optimal testing conditions should maximize the degree to which participants engage fully in the assessment task and respond to all of the items for which they have relevant knowledge. This will minimize response failures currently classified as "not reached" or "off task." There would still be response failures of the type currently classified as "omitted" (students would be instructed to leave an item blank if they had no idea at all about how to respond to it), but these omissions would be desirable (i.e., reflective of a lack of knowledge and thus appropriately treated as incorrect).

Conclusion

Our motivational analysis of educational test taking in general and voluntary twelfth-grade NAEP assessments in particular indicates that NAEP, as conducted in the past, offers nothing of objective value to the participating students, their schools, or their communities. Consequently there are no incentives for participating in NAEP, and some costs, particularly for students with histories of low achievement, test anxiety, or stereotype threat. Consequently, if NAGB's policies require repeating the same procedures that have been used in the past, we would recommend dropping the twelfth-grade assessment from the NAEP program.

However, to the extent that the NAGB's policies would allow for incorporation of the principles and strategies suggested in the section of our report that presents the positive case, there may be reason for continuing or even expanding the NAEP program at the twelfth grade (or at least, conducting pilot efforts to assess the effectiveness of those suggestions viewed as feasible within NAGB guidelines, and then deciding what to do).

We value the opportunity to draw on the literature on motivation in education to provide input to NAGB's deliberations, and we hope that our analyses and suggestions prove useful. We are unable to specify the relative power and effectiveness of the strategies we have suggested, because the literature does not support such specificity and because individual differences among students (e.g., their relative responsiveness to material vs. symbolic incentives or the relative strengths of their orientations toward success seeking vs. failure avoidance) will lead to parallel differences in their responsiveness to different motivational strategies. We can recommend each strategy

with confidence, however, and predict that combinations that include most or all of the strategies will be more effective than combinations that include just one or two of them.

References

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. Journal of Educational Psychology, 84 , 261-271.

Aronson, J., & Steele, C. (2005). Stereotypes and the fragility of academic competence, motivation, and self-concept. In A.J. Elliot & C.S. Dweck (Eds.), Handbook of competence and motivation (pp. 436-456). New York: Guilford Press.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Brophy, J. (2004). Motivating students to learn (2nd Ed.). Mahwah, NJ: Erlbaum.

Cohen, G., Steele, C., & Ross, L. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. Personality and Social Psychology Bulletin, 25, 1302-1318.

Croizet, J, Desert, M., Dutrevis, & Leyens, J. (2001). Stereotype threat, social class, gender, and academic underachievement: When our reputation catches up to us and takes over. Social Psychology of Education, 4, 295-310.

Deci, E., & Ryan, R. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum.

Deci, E., & Ryan, R. (1994). Promoting self-determined education. Scandinavian Journal of Educational Research, 38, 3-14.

Deci, E., & Ryan, R. (Eds.). (2002). Handbook of self-determination research. Rochester, NY: University of Rochester Press.

Eccles, J., & Wigfield, A. (1985). Teacher expectations and student motivation. In J. Dusek (Ed.), Teacher expectancies (pp. 185-226). Hillsdale, NJ: Erlbaum.

Feather, N. (Ed.). (1982). Expectations and actions. Hillsdale, NJ: Erlbaum.

Harackiewicz, J., Barron, K, Pintrich, P., Elliot, A., & Thrash, T. (2002). Revision of achievement goal theory: Necessary and illuminating. Journal of Educational Psychology, 94, 638-645.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. Review of Educational Research, 58, 47-77.

Hill, K., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it. Elementary School Journal, 85, 105-126.

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. Psychological Science, 16, 175-179.

Kaplan, A., & Middleton, M. (2002). Should childhood be a journey or a race? Response to Harackiewicz et al. (2002). Journal of Educational Psychology, 94, 646-648.

Maehr, M., & Midgley, C. (1996). Transforming school cultures. Boulder, CO: Westview Press.

Neveh-Benjamin, M. (1991). A comparison of training programs intended for different types of test-anxious students: Further support for an information-processing model. Journal of Educational Psychology, 83, 134-139.

Pekrun, R. (1993). Facets of adolescents' academic motivation: A longitudinal expectancy-value approach. In P. Pintrich & M. Maehr (Eds.), Advances in motivation and achievement (Vol. 8, pp. 139-189). Greenwich, CT: JAI.

Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. Contemporary Educational Psychology, 25, 68-81.

Zeidner, M. (1998). Test anxiety: The state of the art. New York: Plenum.