

“Flow” as a Testing Ideal

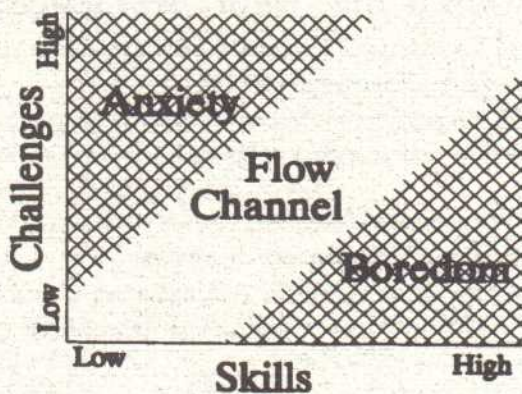
“In our studies, we found that every flow activity... provides a sense of discovery, a creative feeling of transporting the person into a new reality.”

(Mihaly Csikszentmihalyi, *The Psychology of Optimal Experience*, Harper & Row, 1990 p.74)

Csikszentmihalyi describes how human activities often comprise two opposing components, which in the Diagram are characterized as Challenges and Skills. So long as the level of challenge facing the player of a game is in rough accord with the level of the player's skill, then the player will experience a “sense of discovery,” or even a “previously undreamed-of state of consciousness” — that is *flow*. But as the player's skill increases, the player will grow bored. Or when the challenge of the game increases too far beyond the player's skill, frustration will set in. Both boredom and frustration inhibit the flow experience. The motivation towards enjoyment provokes one to desire to balance challenge with skill, and so to induce flow.

Tailored testing can take advantage of the phenomenon of flow to make the testing experience pleasurable and to improve individual performance. Well-targeted items will make the testing situation less irksome, perhaps even enjoyable! Targeting removes items that are too hard, so inducing anxiety, and those that are too easy, so inducing boredom. Psychometrically, the better the match between the item's difficulty and the test-taker's ability, the greater the likelihood that the situation will produce accurate measures. After a test that successfully matches item difficulties with test-taker ability, test-takers can leave feeling content that their optimum performance levels have been demonstrated, and test constructors can count on accurate measures. A flow experience for all!

Craig Deville, Ohio State University



A Savvy Test-taker.

Thomas O'Neil

In 1996, I decided to sit for the GRE. Working for a company that creates and administers several computer-adaptive tests, I knew adaptive tests were usually less grueling than the pencil and paper variety. Being human, I also wondered if there was some way to exploit the adaptive nature of the test. I read in the ETS brochure that there is a minimum number of questions that had to be answered in order to receive a score. This seemed to me to be an opportunity to control when the test stopped. I decided that I would do my best on each subtest up to the minimum number of questions. Thereafter, I would take my time, endeavoring to answer all the questions correctly, but if I was uncertain of the answer, I would stop and wait for time to run out. When I actually took the test, I found myself faced with this situation only once.

I thought that this strategy would help me because all adaptive tests are based on some type of latent trait theory. Usually, the ability estimate is the log of an odds ratio (right over wrong). My thinking was that I wanted to increase the numerator (number correct) without increasing the denominator (number wrong). Provided there was no penalty after answering the designated minimum number of items and provided that I could accurately predict whether I answered a question right or wrong, this strategy should have given me a slight advantage.

As expected, I did reasonably well on the test, but I still wonder what difference the test-taking strategy really made, if any. I have run some simulations using other data sets and found some slight mean increases, but I still don't know what difference it made for me. The experience caused me to think about the role of time in a test and the treatment of incomplete test records in a variety of situations. There has been a great deal of concern about the opportunity to “cheat” on adaptive tests by manipulating responses so the algorithm is more favorable to the examinee. Having attempted a test-taking strategy that was supposed to produce more favorable results, I still have no way of knowing if it worked, or if I would have earned the same score having taken the rest of the test. Obviously, I'll never know, unless ETS asks me to take a complete test, but indicate where I would stop using this strategy and then compare the two scores. Anyhow, I prefer to think that my score is a reflection of my scholastic achievement rather than my ability to devise a “cunning” strategy.

TESTING - TESTING - TESTING

