Reliability and Validity of The Attentional and Interpersonal Style (TAIS) Inventory Concentration Scales

Robert M. Nideffer, Ph.D.

In 1976 I published a book titled, The Inner Athlete dealing with Sport Psychology (Nideffer, 1976a), and an article in the Journal of Personality and Social Psychology (Nideffer, 1976b) on the development of a psychological test called The Attentional and Interpersonal Style (TAIS) Inventory. The Inner Athlete dealt with the role concentration plays in performance and presented a set of theoretical constructs that began to behaviorally define what it means to concentrate. TAIS was a 144 item, self-report, paper and pencil inventory, that had been developed to measure the behaviorally defined concentration styles, and interpersonal characteristics that the theory of attentional and interpersonal style hypothesized were the building blocks of performance.

It is important to point out that TAIS was not developed as a sport psychology measure, but rather as a more general measure of performance relevant characteristics. The goal behind the development of TAIS was to create a tool that would identify an individual’s concentration and interpersonal strengths and weaknesses. A tool that would help identify behavioral predispositions, and in so doing, lead to the development of programs designed to help people improve their level of performance by either strengthening, or better managing performance relevant behaviors in what ever performance arena was important to them.

Both The Inner Athlete, and TAIS, had an immediate impact on the field of sport psychology (Salmela, 1981; Snyder & Abernethy, 1992). What was of most interest within sporting environment were those scales on TAIS, and those aspects of the theory, that dealt with concentration skills. The importance of concentration, the need to better define it, and the need to provide ways to help individuals improve was intuitively obvious to every coach, athlete, and psychological practioneer who had ever been involved in sport. TAIS items and TAIS theory found their way in the Coaching Association of Canada’s training and certification manuals for all their coaches at every level from youth sport to the Olympics (Coaching Association of Canada, 1981). The theory and test were incorporated into the development of athletes at the Australian Institute for Sport from the time in opened in the early 80’s (Bond & Sargent, 1995). Both are being used at Olympic training centers in other parts of the world including Spain (Nideffer, 1989), Italy (Selder, 1982; Cei, et. al., 1997). The Inner Athlete and TAIS were translated into Russian and Chinese and made it behind the Iron Curtain before it came down.

The interest in TAIS and in the theory weren’t confined to the area of application. According to Fogarty (1995), the theory related to attentional styles was the third most heavily researched area within sport psychology between 1974, and 1992. Some of that research raised questions about the reliability and validity of one of the attentional constructs underlying the development of TAIS attentional scales (Van Schoyck & Grasha, 1981; Vallerand, 1983; Albrecht & Feltz, 1987; Dewey, Brawley, & Allard,
1989; Summers & Ford, 1990; Summers, Miller, & Ford, 1991). Before getting into the specific questions raised by researchers, we need to review the constructs that TAIS is based on.

**TAIS Attentional Constructs**

- The intersection of the dimension of width and the dimension of direction of concentration leads to the identification of four distinct attentional styles. A broad-external style used to rapidly assess the environment. A broad internal style used for problem solving and strategic thinking and planning. A narrow-internal style used to rehearse, systematize, and organize information. A narrow-external focus used for execution once an external performance relevant target has been identified.

- Different performance situations require different amounts of shifting and/or place greater emphasis on particular styles.

- Although the average individual has the ability to shift concentration along both dimensions the majority of the time, different individuals have different preferred or dominant concentration styles.

- Under pressure, an individual’s more highly developed or dominant concentration style begins to control that person, rather than the other way around. If the performance situation is one that plays to the person’s dominant style, they perform well, if not, they lose their ability to make the adjustments needed in their focus of concentration to pick up performance relevant cues and begin to make performance and/or decision making errors.

- The specific type of mistakes a person is most likely to make in what is for them a high-pressure situation, are mistakes related to their particular dominant concentration style.

- Choking, or a downward performance spiral occurs when an individual’s level of emotional arousal reaches the point that attention begins to involuntarily narrow, and become increasingly focused on task irrelevant internal thoughts and feelings. This is most likely to happen in those situations where the performance outcome is critical to the individual, where their particular strengths have interfered with their ability to perform (e.g., because the situation required a different style), and as a result of the mistakes, the individual has lost confidence in their ability to perform. It is this sequence of events that leads to the loss of emotional control resulting in what coaches call “choking.”

When TAIS was developed, items were written to assess the ability to develop all four attentional styles. Items were also written to identify the types of mistakes, or the failure to make appropriate shifts in concentration. Subsequent item analyses revealed that items measuring the ability to narrow one’s focus of concentration along both the external and
internal dimensions were so highly inter-correlated that the items were combined into a single scale measuring focus. Likewise, correlations between items measuring errors because of a failure to shift from a narrow external to a narrow internal focus or vice versa, were combined into a single scale measuring errors of under-inclusion. This resulted in the development of six attentional scales:

- BET, a measure of the ability to develop a broad external focus of concentration.
- BIT, a measure of the ability to develop a broad-internal focus of concentration.
- NAR, a measure of the ability to develop a narrow focus of concentration.
- OET, a measure of the tendency toward environmental distractibility or external overload.
- OIT, a measure of the tendency toward internal overload or distractibility.
- RED, a measure of a tendency toward a reduced focus of concentration, resulting in errors of under-inclusion.

**TAIS Attentional Scale Reliability and Validity**

There were two questions being raised by researchers. One, had to do with whether or not TAIS scales designed to differentiate between an internal, and an external attentional style were reliable and/or valid? The other, not unrelated to that question had to do with whether or not the attentional style scales on TAIS could reliability differentiate between levels of performance within sport.

Questions about the reliability and validity of the TAIS construct relating to the existence and measurement of a separate external and internal focus of concentration came from studies that used factor analysis of subjects’ scores on the six TAIS attentional scales to see if the factor structure replicated the six TAIS scales (Van Schoyck & Grasha, 1981; Landers, 1982; Vallerand, 1983; Ford & Summers, 1992.). This research was predicated on the belief that the attentional constructs being measured by TAIS should be uncorrelated with each other, or statistically independent.

What the investigators found, and what I have found in countless studies with every kind of subject population from athletes, to CEO’s, to special operations forces, is that factor analysis of the six TAIS attentional scales does not result in six factors. Instead, it results in two or three factors depending on the analysis, and/or the subject population being studied. The two factors that show up in virtually every analysis and account for about 70% of the variance are a factor reflecting the ability to narrow attention and avoid becoming distracted (NAR, OET, OIT), and a factor reflecting the ability to develop a broad focus of concentration avoiding the tendency to make mistakes of under-inclusion (BET, BIT, RED). The three factor solution accounts for 85% of the variance, pulling the RED scale away from factor two and combining it with the NAR scale which loads on both factor 1 and factor 3. This factor is more common with athletes and high scores on the factor suggest considerable attention to detail and a perfectionistic attitude on the part of the athlete. The conclusions drawn by these researchers were that there was support for the attentional construct of width of attention (broad to narrow focus), but not
for the construct of direction of focus, internal to external (Abernethy, Summers, & Ford, 1998).

As I pointed out in an article in The Sport Psychologist (Nideffer, 1990), the assumption of statistical independence of the attentional constructs, or of any performance relevant cognitive or personality characteristics is inappropriate. Using the attentional constructs as an example, the effective analysis of a problem (BIT) is dependent at least in part upon the individual’s ability to gather or assess relevant external data (BET), and also dependent upon the individual’s ability to organize that data and drive to a solution (NAR). The mutual interdependence of these attentional styles suggests they should be correlated with each other, and if they are, then factor analysis will not identify six factors reflecting the six different scales. Factor analysis will cluster scales together into factors around the constructs that seem to be accounting for the most variance. In this case, the variables accounting for the greatest amount of variance were those reflecting width of focus of concentration, and effective and ineffective processing, supporting these constructs, but because of the limitations of factor analysis could not be used to rule out those constructs not supported.

The position I took in the 1990 article about the inter-relationship of performance relevant variable and the fact that factor analysis of scales is not a reasonable way to attempt to assess the construct validity of an instrument is not unique to me. Indeed when one factor analyzes the three validity scales on the Minnesota Multiphasic Personality Inventory along with the ten clinical scales, they end up with two factors, not thirteen (Dahlstrom, et. al 1975). Factor analysis of the eleven subtests on the Wechsler Adult Intelligence Scale (WAIS) results in between 2 and four factors depending on the analysis (Matarazzo, 1972). The fact that factor analysis of these tests subscales has not resulted in factors supporting their independence, has not been seen as a reason to question the validity of the constructs the various scales were designed to measure, nor has it been seen as a reason to question whether or not the scales measure what they say they measure. Instead, investigators realize they have to find ways other than through the use of factor analysis, to assess the validity of those constructs.

Interestingly, Factor Analysis has been the tool of choice to demonstrate the interconnectedness of various personality and cognitive behaviors. Indeed, the definition of factor analysis provided in the American Psychological Association’s Standards for educational and psychological testing is (APA, 1985): “Any of several methods of analyzing the inter-correlations or co-variances among variables by constructing hypothetical factors, which are fewer in number than the original variables. It indicates how much of the variation in each original measure can be accounted for by each of the hypothetical factors (pg. 91).”

The fact that various intelligence subscales cluster together is seen by some as support for Spearman’s theory of g, or general intelligence, and it is the inter-relatedness of personality characteristics as demonstrated by factor analysis that has led to the “Big Five” personality theory (John, 1990; Digman, 1997). What Factor Analysis of a broad-range of personality inventories has shown, is that the literally hundreds of personality
characteristics which make sense to us conceptually, and that we use to describe the behavior of others, typically cluster together into five broad personality factors (Barrack, & Mount, 1991; Hough, 1992; Tett, Jackson, & Rothstein, 1991; Hogan, Hogan, and Roberts 1996).

The five broad or higher order factors identified by researchers include: 1) A leadership factor; 2) A factor reflecting emotional stability; 3) A factor reflecting conscientiousness; 4) A factor reflecting agreeableness, and; 5) A factor reflecting “intellectance,” or being open, imaginative, broad minded, curious. Not surprising, when all twenty scales on TAIS are factor analyzed they typically result in five factors that to one degree or another can be related to the “big five” factors, and account for between 68 and 80% of the variance in subject’s test scores (Nideffer, 1990).

Does the fact that personality characteristics correlate with each other mean we have to abandon the use of the individual scales that make up those clusters, that they aren’t valid and/or useful? No, in fact just the reverse (Hogan, Curphy, & Hogan, 1994).

“We recommend selecting personality predictors on the basis of job analysis results because measures chosen in this way have significantly higher correlation’s with performance. Next, we recommend matching measures and criteria in terms of their specificity. Although the big-five dimensions are useful for summarizing results, they are the wrong band width for many prediction problems; narrower measures of personality often yield higher validity coefficients.”

What Hogan is saying is although different personality characteristics share common variance with one or more of the five broad factors they are each carving out a smaller portion of that larger factor, a portion that has relevance to a particular type of performance and/or situation. In other words, just because a personality characteristic is not statistically independent doesn’t necessarily mean it is in valid or irrelevant. In fact, that characteristic may be much better as a predictor in certain situations than the broader behavioral category that it is associated with.

Reliability of TAIS Attentional Scales

The same group of researchers using factor analysis to examine TAIS attentional scales questioned were the ones raising questions about the internal consistency or reliability of the attentional scales (Van Schoyck & Grasha, 1981; Landers, 1982; Vallerand, 1983; Ford & Summers, 1992; Abernethy, Summers, & Ford, 1998). Internal consistency is a measure of the extent to which the items within a scale are measuring the same construct. The internal consistency or reliability coefficients found by these investigators for the six attentional scales averaged 64.5 and ranged from .57 to .72. These scores are statistically significant and do indicate the scales have some reliability, not enough, however, in the eyes of this group of researchers.
What is an appropriate level of reliability? Nunnally (1967), in his classic text on Psychometric Theory stated, “In the early stages of research one saves time and energy by working with instruments that have only modest reliability for which purposes reliabilities of .6 or .5 will suffice…. For basic research it can be argued that increasing reliability beyond .8 is often wasteful (pg. 226).”

If one examines the internal reliability coefficients for scales on the Wechsler Adult Intelligence Scale, and on the Minnesota Multiphasic Personality Inventory, they find a number of the scales on both instruments have reliability coefficients below .7 (Matarazzo, 1971; Dahlstrom, Welsh, and Dahlstrom, 1975).

Internal consistency of the items within a scale is going to vary for several reasons. It may vary because the items are measuring different constructs. It may vary because items are measuring more than one construct. It may vary because items are measuring different parts of a broader construct. The more narrowly defined the construct, the more easily understood the items, the less items range in difficulty, and the more situation specific the focus of the items the higher you would expect the internal reliability coefficients to be. Investigators questioning the reliability of TAIS attentional scales were invested in the development of sport specific versions of the TAIS attentional scales and were contrasting the reliability coefficients for those versions with the reliability coefficients they found for TAIS scales.

TAIS was not designed to be sport specific or situation specific. The goal was to create an inventory that would be useful in predicting performance and in providing feedback about concentration skills to individuals involved in a wide variety of performance arenas. To this end, items within scales although created to measure a specific construct like width of attention, differed from each other with respect to the performance arena they focused on. These differences act, as we would expect to reduce reliability coefficients.

The reason reliability coefficients are reduced is relatively simple. The extent to which a concentration skill or a personality characteristic is trait like or state like (changing with situations) varies from characteristic to characteristic and from person to person. The more situation specific a characteristic is, the lower the reliability coefficients we can expect for a general measure, and the higher the reliability coefficients we can expect for a situation or sport specific measure. TAIS internal consistency reliabilities can probably be improved, but they are within an acceptable range provided the scales can be shown to have some construct validity.

**TAIS Predictive Validity**

According to Moran (1996), studies of the predictive validity of have had mixed, and at times confusing results. High scores on the scale measuring a narrow focus, and correspondingly low scores on the external distractibility and internal distractibility scales have been found to predict performance in diving (Nideffer, 1987), and in rifle shooting (Landers, Furst, & Daniels, 1981). Those findings would be expected. The finding by
Wilson, Ainsworth, & Bird (1985) that good concentrators in volleyball to score higher on the narrow focus scale and lower on the scales measuring a broad-external focus, and a broad-internal focus is not so intuitively obvious. Nor is the finding by Kirshenbaum and Bale (1984) who discovered that expert golfers scored higher on both the narrow focus scale and the TAIS scale measuring errors of under-inclusion than novice golfers. Finally, Summers, Miller, and Ford (1991) found that TAIS attentional scales did not differentiate between athletes in the sports of cricket, basketball, and fencing.

For a long time, I was as confused about the results found by Wilson et. al., and by Kirshenbaum and Bale, as everyone else. Nor could I explain the failure of Summers et. al. (1991) to find differences between athletes as a function of level of skill other than to speculate about sample differences, numbers of subjects, levels of expertise, response set issues, and a hundred other possible sources of error variance. Over the past few years, however, through additional research, I believe I have gained some insight that has helped me refine my theoretical constructs, and at the same time explain some of the conflicting findings. In what follows, I am going to first comment on some of the issues affecting the ability of researchers ability to find predicted relationships, and then review the construct and predictive validity of TAIS attentional scales in more detail. I will be adding new data and tying findings to the theoretical constructs presented earlier in this paper.

**Issues In Applied Research**

In my 1987 article in response to questions about the predictive validity of TAIS, I identified a number of factors that make applied research, especially when that research attempts to look at between subject, or between group differences, difficult. I also offered some suggestions for reducing some of the error variance that interferes with the ability to find predicted relationships (Nideffer, 1987). It is worth reviewing a couple of those here as they are related to the data that follows.

One problem researchers face involves being able to get access to enough subjects that are capable of performing at elite levels to be able to design well-controlled studies. One of the advantages I have over many other researchers, as the developer and distributor of TAIS, is access to a tremendous amount of data. That includes data on 239 world champion athletes, over a thousand corporate presidents and CEO’s, all of the data from the Australian Institute for Sport from 1982 through 2005, as well as data from other Olympic training centers around the world.

Perhaps even more important than the numbers of elite level performers one has access to, however, are the conditions under which testing takes place. When data is collected strictly for research purposes, and/or to make selection decisions it affects the willingness of some subjects to cooperate, as well as the honesty of responses. This introduces considerable error variance into the data, dramatically increases the number of subjects one needs to test for significance, and makes between subject comparisons tenuous at best. Once again I am fortunate, in that I have access to data was not collected for research purposes. All of the athlete and business data I will be talking about, including
the data from the AIS, data on world champion athletes, and data on CEO’s of corporations throughout North America, was gathered to be used to provide feedback designed to help improve the performance of those individual taking the test. Because testing conditions were non-threatening, and because the motivation to improve was extremely high among these individuals, responses tended to be more open and honest leading to less “error variance” in the data.

Construct Validity of TAIS Attentional Scales

What evidence is there to show: 1) that different individuals have different concentration strengths; 2) that concentration skills have both state and trait components; 3) that different performance situations require different concentration skills; 4) that individuals gravitate towards and perform better in situations where the concentration demands match their particular skills, and 5) that the type of concentration error a person is most likely to make in pressure situations is directly related to their concentration strength?

When elite athlete’s scores on TAIS concentration skills are standardized based on the norms for the general population, athletes independent of gender, age, or level of performance are dominated by a narrow focus of attention. Not only that, but a comparison of elite level performers at the AIS in Australia with a group of 142 athletes from around the world who had either won a world championship, or an Olympic medal revealed that as the level of a performance for elite level athletes increases scores on the narrow focus scale on TAIS increase, and scores on the broad-internal or analytical scale on TAIS decrease (Nideffer, Sagal, Lowry, and Bond, 2001; Nideffer & Bond 2003).

To investigate the phenomena further, we collected data on an additional 97 world champions bringing the total population to 239 athletes. Next, we divided them into two groups, a group that had won a single world championship or Olympic medal (N=152) and a group that had won multiple medals (N=87). Subjects scores on the BET, BIT, and NAR scales were converted to standard scores and then percentiles based on the standard norms for the test. Thus the percentile scores shown reflect how the groups compare to the general population (Nideffer, Bond, Cei, & Manili, 2003). Figure 1 shows the results of that analysis, and has been included so you can see the pattern of attentional scores found across all elite athlete groups independent of gender, culture, type of sport, or age, when scores are standardized based on the general population.

As you can see in Figure 1, for both groups, the dominant concentration style is focused. A Newman Keuls analysis of the significant interaction (F = 7.38, df = 2, 474, p<.001) shows that multiple medal winners are significantly more focused that single medal winners (p<.01), and significantly less analytical (p=.038). These results are consistent with those of Wilson, et. al., (1985), in that they found elite performers to have lower scores on the BET and BIT scores and higher scores on the NAR or focus scale. The results provide strong support for TAIS’s ability to differentiate between athletes of different skill levels.
Although the results provide support for the ability of TAIS attentional scales to differentiate between athletes as a function of level of performance, they seem to be inconsistent with the notion that different types of sports or performance situations require different concentration strengths, and that athletes gravitate toward and/or perform better in those sports that play to their particular strengths.

**Concentration Skills As A Function of Type of Sport**

When athlete’s scores are standardized based on the general population their high scores on the NAR scale make it almost impossible to find differences in their concentration style as a function of sport type. By standardizing their scores based on their own means and standard deviations you remove the bias created by their exceptionally high NAR scores. When the “average” elite athlete scores at the 50th percentile on each of the attentional scales can see that athletes in different types of sports have different concentration styles (Bond & Nideffer, 1992; Nideffer, 1993).

Recently, I went back to data collected on 4766 athletes at the AIS. There were 2535 athletes involved in team sports like hockey, soccer, baseball, 767 athletes involved in closed skill sports like diving, archery, shooting, and 1464 athletes involved in open skill sports like tennis, judo, and karate. Figure 2 shows the highly significant sport type by concentration skills interaction (F=11.25, df=4, 9526, p<.001).
In looking at figure 2, keep in mind that athletes are now being compared to each other, not to the standard norm group. A Newman Keuls analysis revealed closed skill sports participants continue to be dominated, as would be predicted, by a narrow focus of concentration, scoring significantly lower in both the aware and analytical areas (p<0.001), and significantly higher than the other two groups on the focus scale (p<0.001). Athletes participating in team sports and open skill sports score significantly higher on the awareness scale than athletes in closed skill sports (P<0.001). The difference on the awareness scale between open skill and team sports is significant (p=0.03). The difference between all three groups on the analytical scale is significant (p<0.01). These results support the hypothesis that different sports require different types of concentration, and that athletes gravitate towards and perform better in sports that play to their particular concentration styles. These findings are not limited to sport.

Figure 2

Figure 3 shows the concentration scores of four different groups. A group of 136 CEO’s whose companies were selected for inclusion into the INC 500 list based on their average growth over the past five years. A group of CEO’s who where members of the Young Presidents Organization (YPO). The N for this group was 860. The third group was a business norm group composed of front line and mid level managers (N=1759) and finally, the world champion group (N=239). Scores were standardized based on the means and standard deviations for the general population because the interest was in seeing how these extreme groups in terms of their level of performance performed both in relationship to each, but also in relationship to the standard norm group.

The groups by concentration skills interaction show in Figure 3 was once again highly significant (F=59.47, df=6, 5980, p<.001). A Newman Keuls analysis of the interaction revealed the following. All of the differences shown for scores on the analytical scale are
significant beyond the .001 level. Both CEO groups are more aware than the other two
groups (p<.01). All three-business groups score significantly higher on the analytical
scale than they do on the other two attentional scales (p<.01), and significantly lower on
the focused scale than world champions (p<.001). Finally, world champions are
dominated by a narrow focus of concentration and score significantly lower on the other
two attentional scales (p<.01).

**Figure 3**

These results provide additional support for the ability of TAIS scales to differentiate
between groups based on level of performance, and type of performance situation. They
provide support for the construct of a broad-internal focus of concentration, and a focused
concentration, and support for the hypothesis that high-level performers perform well
because they find performance arenas that play to their strengths.

**Concentration Errors**

Figure 4 shows how the two world champion groups score on TAIS measures of external
distractibility (OET), internal distractibility or overload (OIT), and the scale measuring
under-inclusive errors that occur when one narrows too much and stops shifting between
an external and an internal focus (NAR). The interaction was significant (F=7.30,
df=2,474, p<.001). Percentile scores compare world champions to the general
population.

What Figure 4 shows, and TAIS theory would predict given the ability to focus is their
greatest strength, is that both world champion groups are significantly more likely to
make errors of under-inclusion, than they are to make either of the other types of errors
In addition, multiple medal winners are more likely to make errors of under-inclusion (p<.01) and less likely to make mistakes because they become internally overloaded than single medal winners (p=.02). Again, these results would be predicted by TAIS theory based on their respective scores on the TAIS scales measuring analytical skill, and focus. The results are also consistent with those found by Kirshenbaum and Bale (1984), with elite golfers.

We might want to ask why multiple medal winners would score higher on the scale measuring under-inclusive errors, than single medal winners. It isn’t surprising that this is the highest error score, that fits with the theory, but why would they be more likely than other elite athletes to be under-inclusive? I believe the answer to that can be found in what the RED scale on TAIS measures.

The scale measuring under-inclusive errors was designed to provide an indication of a breakdown in shifting between an external focus and an internal one. The individual either focuses too narrowly on internal thoughts and issues and ignores what is going on around him or her, or focuses too narrowly on the environment and fails to analyze situations when that is what is called for. Because TAIS was designed to measure general behavioral tendencies rather than situation specific behavior, items ask things like, “I can become so involved in something that I am doing, that I lose all awareness of things going on around me.”
Because of the nature of the items on the scale measuring under-inclusive errors whether or not a high score is seen as negative is at times a value judgment. For example, highly dedicated individuals may not pay attention to things going on around them, but they don’t care because they are focused on what is important to them, not to the other person who can’t get their attention. This fact makes it necessary to interpret an individual’s score on the RED scale on TAIS within the context of their scores on other scales. Individuals who score high on the control scale and the self-confidence scale and also score high on the RED scale are perfectionists and highly focused. They make mistakes because they are focused but they don’t see their behavior as a problem because it’s helping them perform as an athlete. Those individuals who score high on the RED scale and score low on the scale measuring control and self-confidence do see their behavior as a problem for themselves, as well as for others. High scores for Kirshenbaum and Bale’s elite golfers (1984) on both the NAR and RED scales on TAIS on the surface didn’t seem to make sense. When you recognize the importance of narrowing, however, and when you look at RED as a measure of perfectionism for those individuals who describe themselves as being in control of their environment (CON) and as confident (SES), rather than as a measure of mistakes, the data make perfect sense.

Figure 5

Figure 5 shows the significant interaction between type of concentration error and type of sport (F=7.129, df=4,9520, p<.001) when subject’s scores are standardized based on their own means and standard deviations. Athletes involved in closed skill sports are significantly more likely to make errors of under-inclusion than they are to become distracted or overloaded (p<.01), and they are more likely to make errors of under-inclusion than athletes in team sports (p<.01) and open skill sports (p=.05). Athletes in open skill sports scores in the three areas do not differ from each other. Athletes in team sports make fewer mistakes overall than the other two groups (p<.01), and are more
likely to become distracted \((p<.01)\), and/or overloaded \((p<.01)\), than they are to become under-inclusive. Their scores on the distractibility and overload scales did not differ significantly from each other \((p=.2)\).

These results provide additional support for the theoretical construct relating concentration errors to concentration strengths. It is interesting that athletes involved in team sports tend to make fewer errors than the other two groups. That may be due to the fact that teammates can serve as supports and reminders, helping them maintain an appropriate focus of attention.

**Figure 6**

![Graph showing concentration errors as a function of vocation](chart.png)

Figure 6 shows the significant interaction between type of concentration error and vocation for the two groups of CEO’s, the world champions, and the business norm group \((F=71.58, df=6,5980, p<.001)\). For this analysis all subjects’ scores have been standardized based on the general norms. As we would predict, world champions most likely mistake is to make errors of under-inclusion and they are significantly more likely to make this mistake than all the other groups \((p<.001)\), and significantly less likely to become distracted \((p<.01)\). Because analytical skills are the greatest strength for CEO’s we would expect their highest score to be on the scale measuring internal overload and it is not. All three-business groups are more likely to make mistakes because they become externally distracted than they are because they become overloaded \((p<.01)\). They are also significantly more likely to make errors because they become overloaded than they are to make errors of under-inclusion \((p<.01)\).

These results provide partial support for the notion that an individual’s concentration errors will be tied to his or her concentration strengths, but does suggest that the OET and OIT scales on TAIS may not be differentiating between these two types of errors as well.
as we would like, either that, and/or the business world, especially the world of the CEO is so demanding that external distractions are unavoidable. There is some evidence to support this hypothesis.

**Gender Differences**

Another source of construct validity for TAIS comes from gender differences found when males and females are compared on both the tests attentional and interpersonal scales. There is compelling neurobiological evidence to support the observations that from an information-processing standpoint, women appear to have more thought linking capacity than men, whereas men have a greater capacity to inhibit information, to focus and process information in a logical sequential way. Women’s perceptual skills are oriented toward quick intuitive people reading, detecting thoughts and feelings absorbing contextual cues and responding in empathetic ways (Kimura, 2001; Marano, 2003; Cahill, 2005).

Studies that have looked at gender differences with respect to scores on the concentration scales on TAIS consistently find that males score significantly higher on the BIT (Analysis) scale, and on the NAR (focus) scale than women. With respect to the BET scale which measures environmental awareness, the result have been mixed, but there have been consistent differences showing that females are more externally distracted than males. These findings hold across cultures (Schmelzer, 1981), when looking at athletes across age levels and type of sport (Bond & Nideffer, 1992), when looking at differences between husbands and wives (LaMotte, 1981).

What is interesting when you look at the patterns of scores for males and females on the effective attentional scales is that there are some marked differences. For females external awareness is almost always either their highest or second highest score and their score on the analytical scale is almost always their lowest. For males the pattern is the exact opposite. Scores on the analytical scale are almost always either the highest, or the second highest, and their score on external awareness is almost always the lowest.

One of the ways to examine these differences is to create an attentional efficiency score for each of the effective attentional scales, and/or ineffective attentional scales by: 1) Standardizing scores based on the population being tested; 2) Taking each scale score once it has been standardized, doubling it, and then subtracting the other two scales in the group from it, for example, (BET*2)-(BIT+NAR). 3) The resulting efficiency scores are then standardized based on their respective means and standard deviations. It is these scores that are used in the data analysis instead of the regular BET, BIT, NAR scores.

The effect of this alteration in scores is to provide an indication of the relative position of subject’s attentional scales to each other. Thus, instead of just looking at between subject differences, which are much more affected by response styles and response sets, you are looking at within subject differences. It is the relative position of the attentional scales to each other, independent of the absolute elevation of the scale that defines a subject’s concentration strengths and weaknesses.
An analysis of variance was conducted, to look at the groups (CEOs with both groups combined, Business Norm Group, World Champions) by gender (male, female) by concentration efficiency scores. Figure 7 shows the significant gender by concentration efficiency scores interaction \((F=13.89, \text{ df}=2,5930, p<.001)\). It is important to keep in mind that the resulting analysis does not really show how males and females compare to each other with respect to the absolute elevation of their scores on the concentration scales. In other words, just because the analysis scale for males is higher than it is for females doesn’t mean they actually had a higher percentile score on the scale. The data show the strength of the relative positions of the scales for the two different groups. What you are interested in is not the elevation of scores across groups, but the elevation of scores within each group. It is the pattern of each group’s scores that is important.

**Figure 7**

As you can see from Figure 7, females score significantly higher on the scale measuring external awareness, than they do on the scale measuring analytical skill \((p<.001)\) and than they do on the scale measuring focused concentration \((p=.05)\). Thus, their confidence is in their ability to read and react to the environment, and much less so in their ability to analyze and problem solve. Males on the other hand have relatively little confidence in their environmental awareness, and much more confidence in their ability to analyze \((p<.001)\) and in their ability to focus \((P<.001)\).

Figure 8 shows the significant gender, by groups, by concentration efficiency interaction \((F=3.659, \text{ df}=4,5930, p=.005)\). Once again, when looking at the figure, it is important to remember that you cannot with this analysis make comparisons across, or between the
different groups. What matters is each group’s pattern of attention scores as it is that pattern that tells you what they perceive their strengths and weaknesses to be.

Looking first at the male and female CEO groups. What you see is that the Male CEO’s perceive their greatest strength is their analytical skill (<.001), followed by their ability to read and react to the environment, which is significantly higher than their ability to focus (p<.01). For females, their greatest strength seems to be their ability to read and react to the environment, which is significantly higher than their score on focus (p<.001) and marginally higher than their score on analysis (p=.06). It would appear as if women CEO’s rely more on their ability to read people and situations than male CEO’s do.

The basic concentration skills pattern found with female CEO’s is found with females in the business Norm group, though these differences are not statistically significant. What they continue to show, however, is the relative strength of awareness for women. For males in the business norm group, the pattern relative to CEO’s changes slightly in that focus is higher than BET. Although these differences are not statistically significant, this is a pattern found when comparing front line managers with senior management. Front line managers tend to score higher on the focus scale. That would be appropriate since they are less involved in strategic thinking and more involved in implementation.

With the world champion athlete groups, for both males and females it is the analytical area that is the area of relative weakness. For males focus is the preferred style and greatest strength (p<.001). The difference for males between aware and analytical is not significant. For females, focus is the strength relative to analytical (p<.001), and external awareness (p=.04). External awareness is significantly higher for females than analytical (p<.01).
It is analyses like these that show both the state, and trait component of the different concentration strengths. Although there appears to be room to move along the attentional dimensions in response to different environments, the consistency for women of the strength of BET in most settings, and of BIT for men in most settings also supports the concept of a trait component, which may vary in intensity from group to group. An analytical focus is more critical and trait-like in the business environment, and a narrow focus appears to be more critical and trait-like at elite levels of sport. Additional evidence of the importance of being born with a predisposition toward, or more of a trait like component with respect to the ability to focus can be found in some test-retest data gathered at the AIS.

**Changes In Concentration Skills Over Time**

TAIS is administered to athletes at the Australian Institute for Sport, when they first enter the institute. The purpose behind the testing process is to assess existing concentration skills, and to then develop a program working with the Institute coaches to help athletes improve their ability to concentrate and avoid making mental errors. Frequently, athletes are re-tested as a part of their development process to see what changes, if any, have occurred, and/or to provide additional direction. The data that follows was collected on 776 athletes at the AIS. Subjects can be broken down in the following ways. There were 244 females, 90 between the ages of 11 and 16, 68 were either 17 or 18 years old, 52 were between 19 and 24, and 34 were 25 or older. Breaking those individuals up on the basis of their sport, 28 were involved in closed skill sports, 104 in open skill sports, and 112 in team sports. The breakdown for the 532 males was as follows. Seventy-three males were involved in closed skill sports, 190 were involved in open skill sports, and 269 were involved in team sports. Seventy-seven males were between the ages of 11 and 16, 121 were either 17 or 18, 211 were between 19 and 24, and the remaining 157 were 25 or older.

Independent of gender, age, or type of sport, when the athlete’s scores were standardized based on the general norms for TAIS, their dominant concentration style was focus, and scores on that scale were significantly higher than scores on the other concentration scales. The fact that focused attention is a dominant score in elite athletes, from a very early age, suggests that this particular type of concentration is critical to high levels of performance in sport. It also suggests that for athletes, the ability to focus may be more trait like, than state like.

The 776 athletes mentioned above were re-tested some eighteen months after their initial test. An analysis of the amount of change that occurred with respect to their concentration skills, and tendency to make errors over that eighteen-month period is presented in Table 1. What you will notice is that athletes’ scores on the focused scale changed significantly less, than their scores on the other five attentional scales. The relatively small amount of change on the focus scale provides some additional support for the belief that this particular skill, at least for this group, has a large trait component to it.
Table 1

Changes In Athlete Concentration Skills Over Eighteen Months

<table>
<thead>
<tr>
<th>TAIS Scale</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>BET</td>
<td>7.0%</td>
</tr>
<tr>
<td>BIT</td>
<td>8.7%</td>
</tr>
<tr>
<td>NAR</td>
<td>2.9%</td>
</tr>
<tr>
<td>OET</td>
<td>8.6%</td>
</tr>
<tr>
<td>OIT</td>
<td>8.7%</td>
</tr>
<tr>
<td>RED</td>
<td>8.4%</td>
</tr>
</tbody>
</table>

Discussion

When is an instrument, or a characteristic being measured by an instrument considered valid and reliable? What are the criteria we should be using to make that judgment? In this paper I have responded to some of the questions that have been raised about the internal consistency of the items within TAIS attentional scales, and about the construct and predictive validity of the instrument.

I have pointed out that the internal consistency coefficients being questioned are similar to those found with other psychological instruments whose reliability is not being questioned. I have discussed the fact that in the real world, performance relevant psychological constructs are complex in themselves, and mutually interdependent. Highly respected multi-dimensional personality inventories and intelligence tests when factor analyzed result in solutions with far fewer factors than scales. That fact alone is not cause to question the construct validity of the scales.

I have provided additional evidence in this paper for the ability of TAIS attentional scales to differentiate between levels of performance at the highest levels in sport, and in business (Nideffer, et. al., 2001). In the process I have providing additional construct validity for scales measuring focused concentration and analytical skills, and helped clarify earlier research findings that supported the predictive validity of the test, but were confusing based on our knowledge at the time.

I have added data to the earlier findings showing that elite athletes involved in different types of sports had higher scores on those particular concentration scales that according to the theory would be most important for their sport (Nideffer, 1990; Bond & Nideffer, 1992; Nideffer, 1993). I have shown how gender differences with respect to concentration skills are consistent with behavioral observations and neuro-biological research, providing support for the differentiation of external and internal attentional processes.
Will the additional data provided in this paper change the opinions of those researchers and reviewers questioning the validity and reliability of the inventory? I have responded to these same criticisms before, by pointing out the methodological problems associated with factor analysis, and by providing data to support the construct validity of the scales. The arguments and data have been brushed over and largely ignored. After challenging the construct validity and reliability of TAIS attentional scales Moran (1996) says, “but see Nideffer, 1990 for a rebuttal of these criticisms.” Moran does not present the arguments, nor does he debate the arguments, he ignores them. In the review by Abnernethy, Summers, and Ford (1998), two of the articles I published that deal with methodological issues and provide data to support the constructs are cited in the text and bibliography, but none of the arguments or data are presented (Nideffer, 1987; Nideffer, 1990). Questions about the validity and reliability of the inventory are accepted without question and passed on, why? I believe the following quote from Abnernethy, Ford, and Summers (1998), provides a clue to the answer to that question.

“Although the TAIS was designed as both a research tool and feedback device, there is little strong empirical support for its use as a research instrument to examine the relationship between attentional abilities and sport performance. There is, however, some support for its use as a diagnostic tool for helping athletes to identify attentional problems that may be affecting performance.” (pg. 188)

Notice the split between “science,” and “application” contained in that quotation. Does it make sense? How does a test that is useful as a diagnostic tool, for helping athletes identify and work on the attentional problems the instrument was designed to measure, not have validity and reliability?

Both scientists and practitioners have a tendency to become wedded to their tools. So much so at times that our pre-conceived beliefs prevent us from seeing, believing, and even understanding arguments that run contrary to our particular philosophical and/or theoretical beliefs. In my rebuttal to critics, I challenged a methodology they believe in, a methodology that fits with their concept of what scientific research is about. I did not say the data they gathered was inaccurate, that the reliability coefficients they obtained were wrong. Instead, I argued with their interpretation of the meaning of the data. Believing in their methodology and in the need for assessment tools that measure statistically independent constructs they could not accept my arguments. They would agree the tool is useful, and does what it was designed to do, but couldn’t or wouldn’t allow that fact to cause them to call into question either their methodology, or the interpretations they were drawing.

Pure science seeks to increase the understanding, prediction, and control of nature by first, reducing everything to its simplest form, and then slowly recombining the elements under highly controlled conditions, to study their interactions. The goal is to increase our understanding and in the end to create a science that will allow us to make very specific predictions with an extremely high degree of accuracy, thereby increasing our control over ourselves, and the world around us.
Although I believe most of us would applaud the goal of pure science, and see tremendous value in the process, we often have problems with both the speed with which advances are made, as well as with the lack of immediate practical value obtained from that research. In my mind, the goal of applied research is to find a happy medium between basic laboratory research on the one hand, and faith in untested hypotheses about cause-effect relationships that on the surface appear to be important but when closely examined are found to be little more than superstitions.

As a scientist, unconcerned about the practical application of my research, but focused on identifying the purest, most basic aspects of human performance, I would focus as critics of TAIS have, on trying to purify measures, to identify constructs that are statistically independent and to then create measures of those constructs. I would use statistical tools similar to the ones those researchers have used, and I would draw the same conclusions they have drawn from the results they have achieved.

My ultimate goal would be to gain the understanding required to create an instrument that would measure those independent performance relevant constructs with enough accuracy for me to be able to use test information to predict the specific behaviors and performance outcomes that would occur within specific performance settings. I would be trying to create an instrument that would take the human element, and the subjectivity that goes with that, out of the prediction equation.

**Conceptual vs. Statistical Independence**

As an applied researcher, as someone who is concerned both about the practical application of my research, and the validity and reliability of the theoretical constructs and tools that I use, I have different expectations and different goals. First, I am more concerned with the conceptual independence of constructs than I am with their statistical independence.

I know that human performance is complex and that performance outcomes are determined on the basis of the interactions between a broad range of performance relevant characteristics or behaviors. If I want to deal with real life issues I have to accept this fact and for that reason realize there are serious limits to the extent to which I can create measures, and/or design studies where relevant behaviors are completely statistically independent of each other.

Conceptual independence means there is sufficient statistical independence to talk about different constructs as if they were independent, even though correlations do exist between them. As an applied researcher, the key question I have to ask myself is, when the common variance measured by two conceptually independent constructs is extracted, do the constructs prove to be valid and reliable measures of what they were designed to measure?
If I am willing to accept conceptual independence as opposed to statistical independence, then the fact that factor analysis of TAIS scales doesn’t result in the identification of statistically independent factors associated with those scales by itself, doesn’t trouble me. Not as long as the scales are still able to show different and predictable relationships with the conceptually independent behaviors they were designed to assess.

The rule of thumb you apply when evaluating the reliability and validity of a theoretical construct or tool is in part pre-determined on the basis of your position relative to the importance of statistical vs. conceptual independence. So too, your judgment as to the required level of reliability and validity of a test will depend upon your ultimate goal, and expectations for how the instrument will be used.

**Testing vs. Assessment**

Those psychologists who use psychological tests as a part of their practice for purposes of diagnosis, selection, counseling, or development, realize that there is sufficient error in the testing process to preclude the possibility of a test being used to make a decision about a person without that decision being consensually validated through other means (e.g., other tests, behavioral observations, past history, further exploration through the additional testing and/or questioning, etc.). Differences between people on some of the constructs we measure like speed of decision making, the willingness to take risks, the willingness to self-disclose, level of anxiety, and level of self-awareness are all examples of factors that will influence subject’s responses to tests and affect our ability to make accurate between subject comparisons based on test data alone.

For this reason, psychologists do not test, they assess. Testing is an event that results in a set of scores. Assessment is a process that often includes testing as a part of the process. Tests are administered and the results from testing are then used to generate hypotheses not about behavior within specific contexts or time frames, but about behavioral predispositions or tendencies. We don’t expect to be able to predict that person A will become distracted during the Olympic final and therefore lose the competition anymore than a geologist or seismologist expects to be able to predict a 7.0 magnitude earthquake in Southern California tomorrow.

The reason the American Psychological Association and test publishing companies limit the sale of psychological tests to individuals who have had courses in statistics and in psychological testing and assessment is to ensure that they are aware of test limitations, of the need to consensually validate results. There is also the hope, that they will have the skill sets necessary to engage in the consensual validation or entire assessment process.

Obviously, we want to create tests that have as much reliability and validity as we can, and the process of fine-tuning, refining, and improving our tools should be constant. However, when one recognizes that: 1) responsibility for the appropriate use and interpretation of tests reside with the test administrator; 2) test data is used not to make decisions but to generate hypotheses about probably behaviors and/or issues, and; 3)
those behaviors are then consensually validated through the assessment process, tolerance for less than perfect reliability and validity coefficients increases. In fact, I can argue that from a practical perspective, if your validity and reliability coefficients are too high, you aren’t measuring anything of relevance.

The Need For Variability In Scores

Most tests were not developed for research purposes, they were developed to help people. The questions being asked weren’t simple. To answer them you need data from many different areas. Often, the data that provides the greatest insight is the disconnect that occurs between thoughts and feelings, or between peoples perceptions. The integration of that information is where the expertise lies it isn’t found in a computer.

Most psychological tests are not developed for research purposes. Tests are developed to provide answers to questions that are important to people. These are questions that don’t have simple answers, if they did, there would be no need for the test. Especially when working with normal, healthy, and often very high functioning individuals, the questions are not what questions, but instead are why, and how, questions. Coaches, athletes, business executives, don’t need a psychologist to tell them what their mistakes are. More often than not their mistakes are painfully obvious to them, and to everyone else. When they approach a psychologist it is because they have identified the issue but been unable to resolve it and want to know why it continues to be an issue and what to do about it. Psychologists develop tests to get at the root cause of issues and to help provide answers around what to do about those issues.

If you have been in the business of helping people get at the root cause of performance issues for any length of time, you recognize that problems are more complicated than they might seem on the surface. One of the first mistakes individuals made with respect to TAIS, was to ignore the scales measuring interpersonal characteristics and attitudes and focus exclusively on the concentration scales. They didn’t realize that problems with concentration, like problems with anger, or communication are multi-faceted.

When people make mistakes, whether the mistakes manifest themselves as concentration or decision making errors, or as inappropriate emotional expressions, or insensitive interpersonal behaviors, to understand the mistakes you must determine the role concentration plays, the role emotions and interpersonal processes play, and the role the environment plays. To do this you have to assess all of these areas, not just concentration, not just emotions, and not just the situational variables. What you are looking for and what good tests help you find, when mistakes repeat themselves, are interactive patterns, unique to the individual that increase the likelihood of the mistake. The same is true of exceptional performance you look for patterns, for the coming together of concentration skills, personal and interpersonal variables within the context of specific situations that increase the likelihood of success.

You cannot separate concentration from emotions or emotions from concentration. Nor can you separate either of those from the environment. It is patterns of behavior that
predict success, not individual variables. Because the patterns are complicated and it is difficult if not impossible to find enough subjects to examine all of these complex interactions researchers select the one or two variables they hope will have the most predictive power. The end result of ignoring many of the contributing factors is less predictive validity.

It is perhaps not surprising that most of the critics of TAIS have been invested in the development of sport specific measures of the instrument. They have done that for two reasons, first to increase the internal consistency or reliability of the measure, and second to increase it’s ability to make between subject predictions about performance. Is that really what we want to do with testing and assessment?

If your goal is to create a research tool to be used to try and zero in and refine a particular construct then doing everything you can to control subject response sets, and reduce every kind of error variance you can imagine makes sense but that isn’t why most tests are created. Most tests are created for applied reasons, not for research, tests are developed to be used for feedback and evaluation.

Coaches don’t come to psychologists to have them predict performance outcomes. Coaches know the best predictor of outcome is past performance and coaches know the athlete’s performance as well as anyone. When coaches ask psychologists for help it’s because the answers to the questions they have aren’t obvious. They come because in their minds the answers should be easy and they aren’t. They come because they have tried all the obvious things and nothing is working. They come because they don’t understand the whys behind what is going on and/or don’t have the knowledge and skills necessary to figure out what to do about it.

When a coach or a business manager asks a psychologist to assess his or her team what are they asking for? In selection situations they are asking questions like: 1) Will the person fit into the team and/or system that is in place? 2) Does the person have the mental and interpersonal skill sets required by the job? 3) What are the person’s strengths and weaknesses? 4) Under what conditions will she perform well and when will she be likely to make mistakes?

In situations where a problem has been identified or where there is opportunity for development whether the issue is seen as a problem or not, the questions focus on: 1) Getting at the root cause of the issue, and 2) Finding out what to do to change in positive ways. Often associated with these questions is the additional question about the likelihood of change and the cost of change in terms of time, money, and effort.

Answers to the kind of questions that really matter are complicated and require data from several different sources. This is true even when on the surface the questions appear to be fairly straightforward like “Why can’t this person stay focused through an entire game?” or “What do I have to do to keep his anger under control?” To get at the root cause of those issues you have to engage in a complicated assessment process. The administration of a test like TAIS is only one part of that process. In addition to
measuring the concentration and interpersonal building blocks of performance, you need to consensually validate test results by comparing them with behavioral observations and data gathered from an interview with the individual.

**Bibliography**


Cahill, L. (2005), His Brain, Her Brain, Scientific American, August, 2005.


