Op-Ed/Editorials – Going for the Gold A Guest Editorial by Blase Gambino, with comments from Marc Potenza and David Korn

March 2, 2005

The recent commentary exchange on gold standards between Claus and the WAGER staff (2004;
touched on a number of interesting issues. Questions about gold standards or the lack of them, as the WAGER notes, are important and merit discussion. The *WAGER* clarified several important concepts and noted some of the difficulties that prevail in their response to Claus but their focus was on diagnostic criteria and diagnostic instruments. The construct of the gold standard, however, may be viewed within a wider and more general framework. One in which a number of gold standards will be needed.

First, let me clear up a common misunderstanding. The search for a perfect gold standard for gambling disorders is most likely the equivalent of the quest for the Holy Grail; it may or may not exist, and its presumed power is likely exaggerated. Put another way, the theoretical construct of the gold standard, as a test that never produces a misdiagnosis, exists more in the realm of test mythology than in the world of test reality (Kraemer, 1992). This is inherently true for any disorder that resides along a continuum of severity. There will be some level below which the disorder is undetectable. Further, in view of the likely multi-factored nature underlying the causes of most disorders and tossing in the concept of interaction (e.g., gene-environment), even the promise of genetic testing is unlikely to produce a perfect diagnostic system (Botto & Khoury, 2001). In other words uncertainty will most likely remain.

This is not to say that a gold standard cannot be developed; or that one would have little value if it is not perfect. That's another misconception: that a gold standard must be errorless to be accepted as a gold standard. What needs to be recognized is that a diagnostic system may be acceptable as the gold standard despite the presence of error. In truth, the amount of error in current gold

standards can be considerable (Zhou et al, 2002). The concept of the errorless gold standard for gambling disorders is better viewed as an ideal that clinical and other researchers are striving to realize (e.g., Cunningham-Williams & Cottler, 2001). The pragmatic goal of this research is to produce, if feasible, a test that has been validated as so accurate and reliable that it is viewed as providing a definitive diagnosis at an acceptable level of uncertainty. Clearly the question to be resolved is "how much uncertainty is acceptable"?

This leads to the second issue, as with any test, the critical question that needs to be asked is, for what purpose is it to be used? The gold standard is no exception. The choice of a gold standard will depend upon the information required and for what purpose. Is it to make a clinical decision on the true status of the patient or client? Is it to be used to determine a treatment plan? Will it be used to screen a population? Is it to be used for the purposes of research such as in the study of the etiology or natural history of the disorder? Or to predict who will develop the disorder? Or to predict who will seek treatment? Our willingness to accept uncertainty will clearly depend upon the purpose for which the standard will be used.

For example, the randomized clinical trial (RCT) is viewed as the gold standard for the evaluation of treatment outcomes. In this situation the term is being applied to a methodology and not to a diagnostic test. In the sense used by the WAGER staff, the gold standard information being sought was the presence or absence of a gambling disorder. In the RCT, the information required is the relative effectiveness of one treatment versus another or versus a placebo. It is the truth of the information that is being evaluated and the standard for the evaluation of this truth, the gold standard, is the "source of the information" (Zhou et al, 2002, p.15).

Third, it is also important to emphasize that a gold standard does not have to be a single test: another common misconception. It is an information-gathering procedure or process. The gold standard may be a single test, a series of tests, or even a mathematical equation or it may occur in the form of a program that eventually produces the desired result (Berner, 2003). It is the accuracy and usefulness of the information produced for a particular purpose that defines the gold standard, not the specific procedure. In sum, the gold standard is the most accurate system that is currently available for acquiring information useful for some specified purpose that is, diagnosis, evaluation, etc.

A fourth issue to be considered is the relative impact of the gold standard. For example, the availability of a gold standard diagnosis generally has its smallest diagnostic impact in the research setting. The reason is that even when a gold standard is available, a relatively rare occurrence, it is generally too costly (e.g., MRI), too invasive (e.g., biopsy) or too complex, to be of practical use for research; particularly in large-scale epidemiologic studies such as those that estimate the prevalence or incidence of gambling disorders (the work by Cunningham-Williams & Cottler, for example, reflects in part, one effort to overcome this difficulty). These same factors, e.g., cost, mean that gold standards will generally not be useful for the purpose of screening populations for referral to treatment; it simply wouldn't be cost-beneficial, particularly if prevalence is low. In the case of screening the gold standard is more likely to be used to verify the results of the screening but not for the purpose of screening.
In the clinical setting the impact of the gold standard will depend, in part, upon the demonstrated effectiveness of an available treatment. The value of the gold standard under these conditions is that when a beneficial (let's ignore the meaning of beneficial since that would take us in another direction) treatment exists then it may be applied with confidence in the [gold standard] diagnosis and the knowledge that the client will benefit from its application (Shaffer & Gambino, 1990). This is not to say that in the absence of an effective treatment a gold standard has no value. Another factor that determines the value of a gold standard is the seriousness of the disorder. Learning one has a terminal or untreatable condition at least allows you to put your affairs in order, or to get help from a support group.

In the final analysis, any discussion about gold standards should recognize that it is a philosophical as well as a methodological and statistical issue. There are those, for example, who "argue that there is no such entity as a true gold standard" (Zhou et al, p. 71). In practice it is a matter of establishing "operational standards for diagnostic truth" (p. 71). How to do the latter I leave to those more knowledgeable than I (e.g., Streiner, 2003).

Having a gold standard does not guarantee we will obtain the knowledge needed, for example, to reduce or eliminate gambling disorders. The RCT, for example, has been criticized on a number of grounds such as the use of exclusion rules that limit its generalizability (Penston, 2003). On the other hand it is clear that a great deal is being learned despite the absence of a gold standard (e.g., National Research Council, 1999). There are a growing number of intrepid researchers

who, unfazed by the lack of a gold standard, are producing quality research that is both informative and generating valuable knowledge (e.g., Grant & Potenza, 2004); no doubt while being forced to work with silver and bronze standards! A gold standard might indeed be useful, but its absence is not a deterrent to identifying who will or will not become a disordered gambler, how to prevent them before they do, or how to treat them after they have; nor does its presence guarantee its usefulness for all purposes.

References

Berner, E.S. (2003). Diagnostic decision support systems: How to determine the gold standard? Journal of the American Medical Informatics Association, 10, 608-610.

Botto, L.D. & Khoury, M.J. (2001). Commentary: facing the challenge of gene-environment interaction: the two-by-four table and beyond. American Journal of Epidemiology, 153, 1016-1020.

Cunningham-Williams, R.M. & Cottler, L.B. (2001). The epidemiology of pathological gambling. Seminars in Clinical Neuropsychiatry, 6, 155-166.
Grant, J.E. & Potenza, M.N. (2004). Pathological gambling. A clinical guide to treatment. Washington, DC: American Psychiatric Publishing, Inc.

Kraemer, H.C. (1992). Evaluating medical tests: Objective and quantitative guidelines. Newbury Park, Calif: Sage.

National Research Council (1999). Pathological gambling: A critical review. Washington D.C.: National Academy Press.

Penston, J. (2003). Fiction and fantasy in medical research. The large-scale randomized trial. London: The London Press.

Shaffer, H.J. & Gambino, B. (1990). Epilogue: Integrating treatment choices. In H.B. Milkman & L.I. Sederer (Eds.). Treatment choices for alcoholism and substance abuse (pp. 351-375), Lexington MA: Lexington Books.

Streiner, D.L. (2003). Diagnosing tests: Using and misusing diagnostic and screening tests. Journal of Personality Assessment, 81, 209-219.

WAGER (2004). Op-Ed/Editorals: Letters to the Editor. Available at

http://www.thewager.org/Backindex/editorials/editorial10062004.pdf.
Zhou, X., Obuchowski, N.A. & McClish, D.K. (2002). Statistical methods in diagnostic medicine. New York: Wiley Interscience.

Mining for Gold: On Gold Standards, Gambling Behaviors, and Gambling Disorders

An Invited Comment by Marc Potenza, BASIS editorial board member

In Mr. Claus' letter to the editor and the corresponding response from the staff of The WAGER (http://www.thewager.org/Backindex/editorials/editorial10062004.pdf), several important topics are discussed. The first topic, as described by Mr. Claus, is the relationship between "actual gambling behavior (e.g., monetary loss, time spent gambling)" and a "GOLD STANDARD" for a "disordered [gambling] state". As The WAGER response indicates, this relationship is complex and not always easily identifiable. One might consider the American Psychiatric Association's diagnostic criteria for pathological gambling as the current "gold standard" for defining disordered gambling (2000). The current guidelines for diagnosing mental health disorders can be found in the Diagnostic and Statistical Manual, Fourth Edition, Text-Revisited, or DSM-IV-TR (2000). This text contains the current diagnostic criteria for all mental health disorders, including that for pathological gambling. Pathological gambling is currently classified as an impulse control disorder, a central feature of which is the diminished ability "to resist an impulse, drive, or temptation to perform an act [gambling] that is harmful to the person or others" (DSM-IV-TR, 2000). The inclusion criteria for pathological gambling include preoccupation with gambling, gambling-related tolerance and withdrawal symptoms, repeated unsuccessful attempts to cut back or quit, gambling to escape dysphoria or boredom, "chasing" losses, lying about gambling, gambling-related illegal activities and financial "bail-outs", and jeopardizing or losing significant life opportunities because of gambling (DSM-IV, 2000). Missing from the diagnostic criteria are quantity/frequency measures of gambling. One reason why the groups of experts involved in originally crafting and later revising the diagnostic criteria for pathological gambling did not include quantity/frequency measures might involve the variability across individuals in levels of gambling behavior that interfere with functioning in life domains. In clinical practice, I have encountered marked variability in the amount of financial loss or time spent gambling that has significantly impacted individuals' functioning. For example,

gambling losses of $20 per week can result in significant impairment for an individual with chronic mental illness living on a limited budget while losing $100 per week in poker gambling can be of little significance to a business executive with a six-figure annual salary. Thus, quantity/frequency measures alone do not necessarily provide meaningful information about the personal, familial or social impact of one's gambling behaviors. The complex relationship between quantity/frequency measures of gambling and gambling-related symptomatology makes it challenging to generate guidelines for "healthy" levels of gambling based on quantity/frequency measures alone.

A second point involving the development of a "gold standard" for diagnosing pathological gambling is the evolutionary nature of this process. Frequently I used the word "current" in the preceding paragraph to underscore the continuing process of refinements made according to available knowledge. Currently, pathological gambling, like other psychiatric diagnoses, is incompletely understood and its origins likely involve a complex mixture of genetic and environmental influences (Grant & Potenza, 2004). Over time, the DSM has incorporated emerging data to guide the most appropriate definition and categorization of psychiatric disorders. As new techniques (e.g., brain imaging, molecular genetics) provide an increased understanding of the mechanisms involved in the etiologies of specific psychiatric disorders, it is likely that scientists and practitioners will generate new or refined diagnostic criteria or "gold standards".

A third point involves the use of casino databases. I agree with both Mr. Claus and The WAGER staff that large, anonymous databases such as those held by large casino enterprises could provide valuable information regarding the gambling behaviors of patrons to specific casinos. Results from such studies would need to be interpreted cautiously given that gambling behaviors might be limited to those occurring within a single casino (or a limited number of casinos operated by one business entity) and might not contain gambling at other casinos or non-casino venues. Furthermore, the databases likely would not capture the diagnostic or "gold standard" information that is arguably valued most by psychiatrists and other mental health workers. Nonetheless, researchers should be encouraged to pursue such data if they could be used to answer specific research questions that have clinical and societal relevance.

References

American Psychiatric Association Committee on Nomenclature and Statistics. (2000).

Diagnostic and statistical manual of mental disorders (4th Ed. – Text revision). Washington, DC, American Psychiatric Association.

Grant, J. E., & Potenza, M. N. (Eds.) (2004). Pathological gambling: A clinical guide to treatment. Washington, DC: American Psychiatric Press, Inc.

On Going for the Gold by Blase Gambino

An Invited Comment by David Korn, BASIS editorial board member

The ongoing dialogue on the concept, purpose & value of a gold standard in the gambling field is both important and challenging.

Blase Gambino extends the boundaries of the concept beyond its role in diagnosis and measurement by discussing misconceptions, limitations and other potential applications to epidemiological research and clinical evaluation. I would like to extend the dialogue even further by picking up on a quote from Dr. Gambino's piece, as follows: "The construct of the gold standard, however, may be viewed within a wider and more general framework." Importantly, he also briefly points out that gambling behavior exists along a continuum.

An expanded public health framework acknowledges the value and precision of a diagnostic gold standard for gambling disorders. However, it utilizes and builds upon the contrasting and complementary concept of a continuum. This notion of a continuum can be applied to a number of gambling dimensions. It can exist for gambling behaviors (healthy to unhealthy), gambling risks (low, medium, high), and gambling problems (mild, moderate, severe).

The practical importance of the continuum concept relates to its potential application for planning, implementing and evaluating prevention, health promotion, harm reduction and treatment interventions. In addition, it shifts focus from solely clinically diagnosed gambling disorders to a broader view of potential gambling-related harms as well as beneficial consequences.

The BASIS Responds

> *We first want to thank Blase Gambino, Marc Potenza, and David Korn for contributing such a thought provoking editorial forum. The discussion of gold*

standards and what we might expect of our field in terms of diagnostic reliability and validity is important. That said, we define a gold standard as an independently verifiable measurement or assessment. In other words, to achieve gold status, the benchmark of diagnosis must be independent of the disorder of interest. Currently, like many other psychiatric disorders, clinicians and patients alike identify gambling disorders by the presence of adverse consequences. This is not consistent with the scientific method. Independent variables must be identified and determined by factors unrelated to their consequences: anything less represents a tautology.

We should note, however, that while we do argue that gold standards are vital to effective treatment, in contrast to what Gambino suggested, we do not propose that a "perfect" gold standard is necessary for diagnosis and treatment. Perfection is rare and that standard is beyond gold. We require confidence in our gold standards, even though these might not be perfect. We believe that we can do better than the existing tools currently in use by clinicians and researchers who treat and study intemperate gambling. We suggest that emerging neuroscience research is likely to evolve into a gold standard for diagnosis or support the development of such a standard.

This brings us to the final point: we disagree with Potenza when he says, "one might consider the American Psychiatric Association's diagnostic criteria for pathological gambling as the current "gold standard" for defining disordered gambling." We do not agree that DSM is a gold standard – it is merely the best available standard – and some might argue even with that suggestion. We suggest that all diagnostic classification—whether clinician- or instrument-based—be held as tentative, and not the final word. The best available diagnostic tool, is just that, the best available. Standards that are available and widely used do not elevate them to the level of a gold standard—that requires much more.

— The BASIS