

# Comparing alternative diagnostic criteria sets using Item Response Theory (IRT)

Melanie M. Wall and Deborah S. Hasin

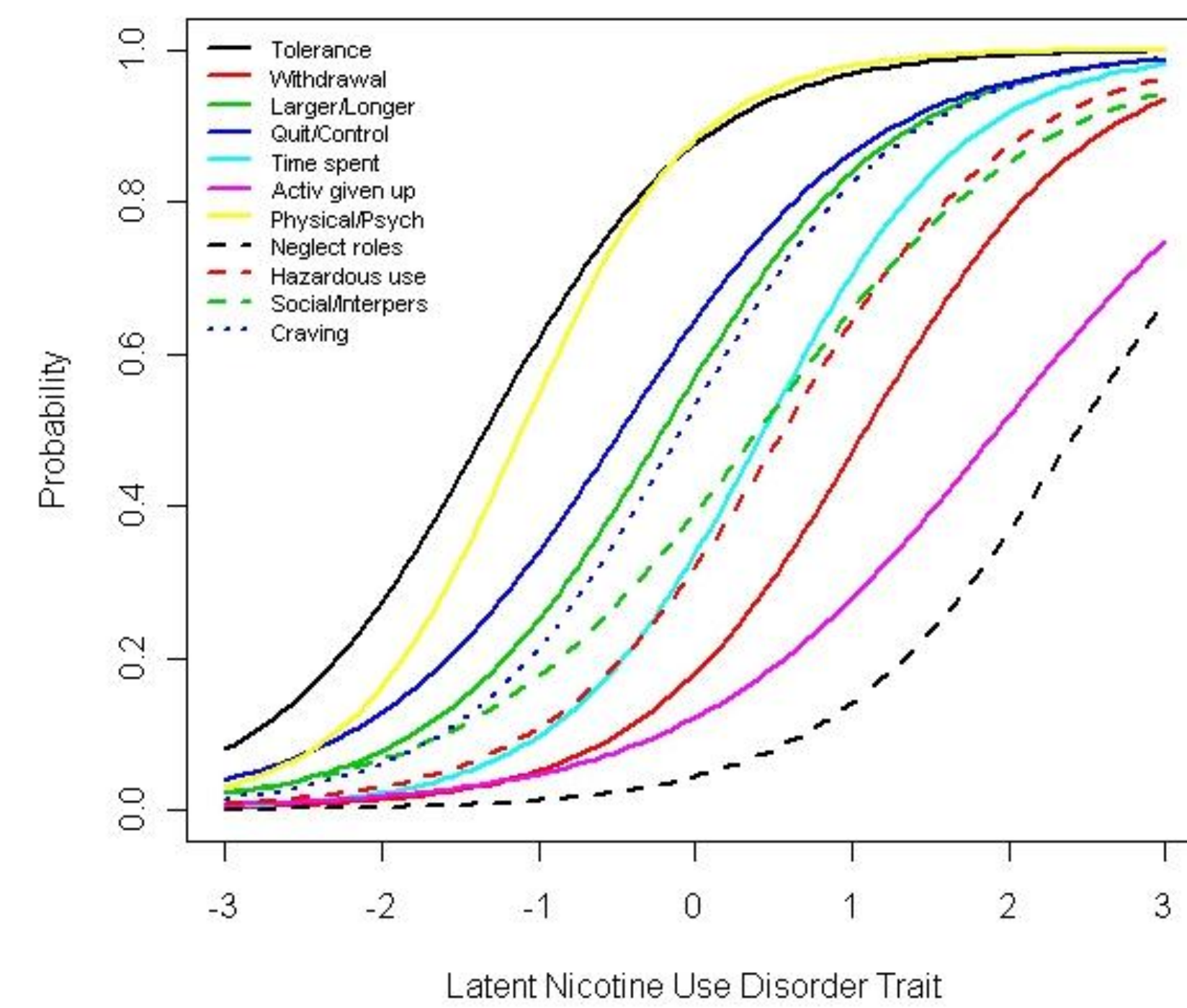
Division of Biostatistics and Substance Dependence Research Group



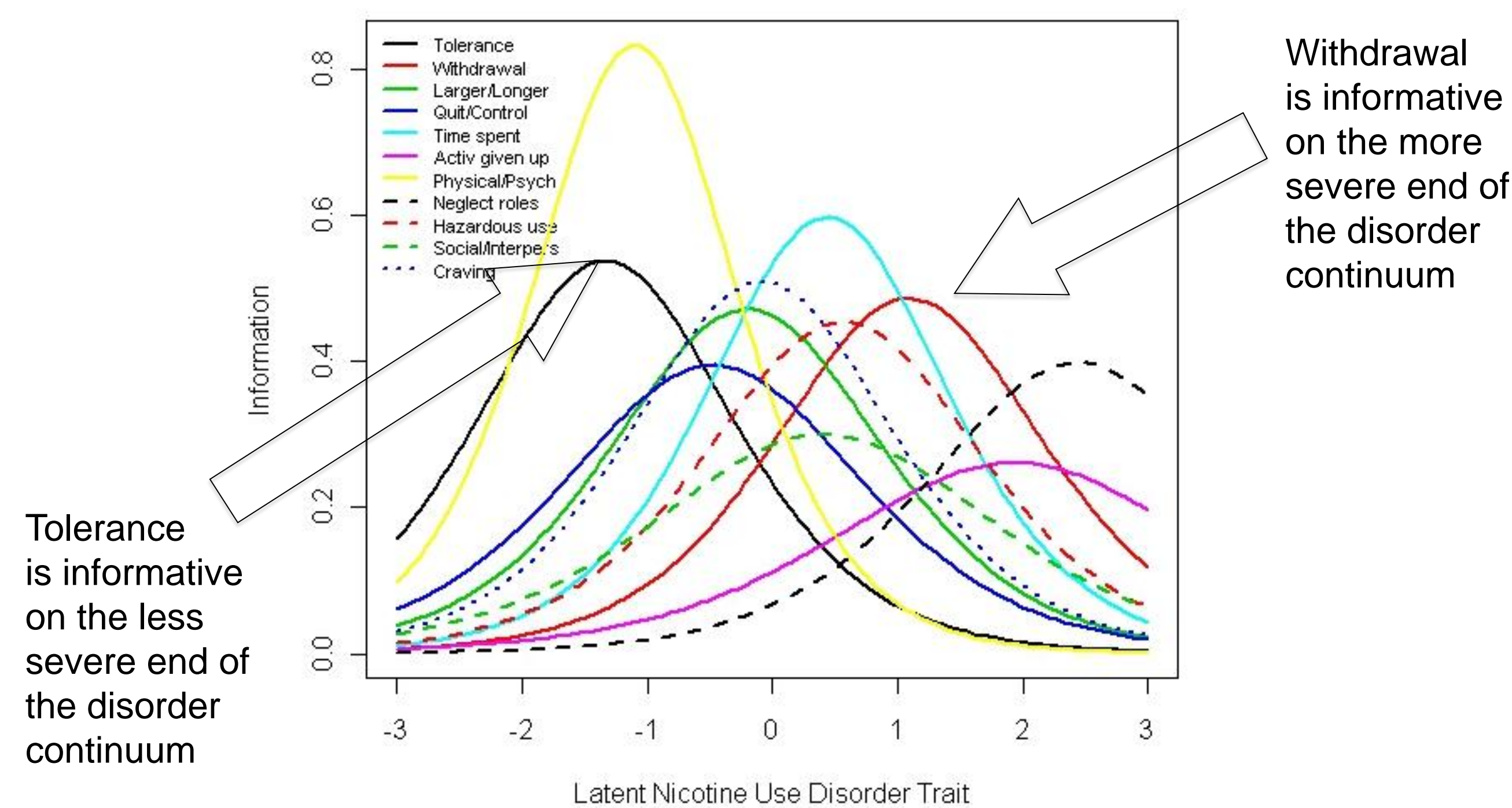
## Item Response Theory (IRT)

IRT models are useful for assessing the ability of criteria sets to measure underlying psychiatric disorders, e.g. substance use disorders. IRT assumes the underlying psychiatric disorder is inherently continuous ranging from low to high severity and estimates how well each observable criterion measures different points along that continuum.

The 2-parameter IRT model is commonly used and describes each of the diagnostic criterion in terms of its **discrimination** (slope) and **severity** (location) for measuring the underlying disorder



**Figure 1.** Curves show the probability (estimated with IRT) that a person will exhibit a particular nicotine disorder criterion as a function of their location on the underlying disorder continuum. A person's true location on the underlying disorder continuum (x-axis) is not observed directly but can be inferred from the IRT based on their particular combination of observed criterion.



**Figure 2.** Criterion information curves represent the precision ("information") with which each criterion measures the underlying psychiatric disorder across the continuum.

## Nicotine Use Disorder

We illustrate IRT with an application to Nicotine Use Disorder. With the changes proposed for the other DSM-5 SUDs to combine dependence and abuse criteria along with cravings into a single disorder, we examined whether combining nicotine dependence, abuse and cravings criteria is warranted.

**Can the measurement of nicotine use disorder be improved by adding nicotine abuse and cravings criterion to existing nicotine dependence criterion?**

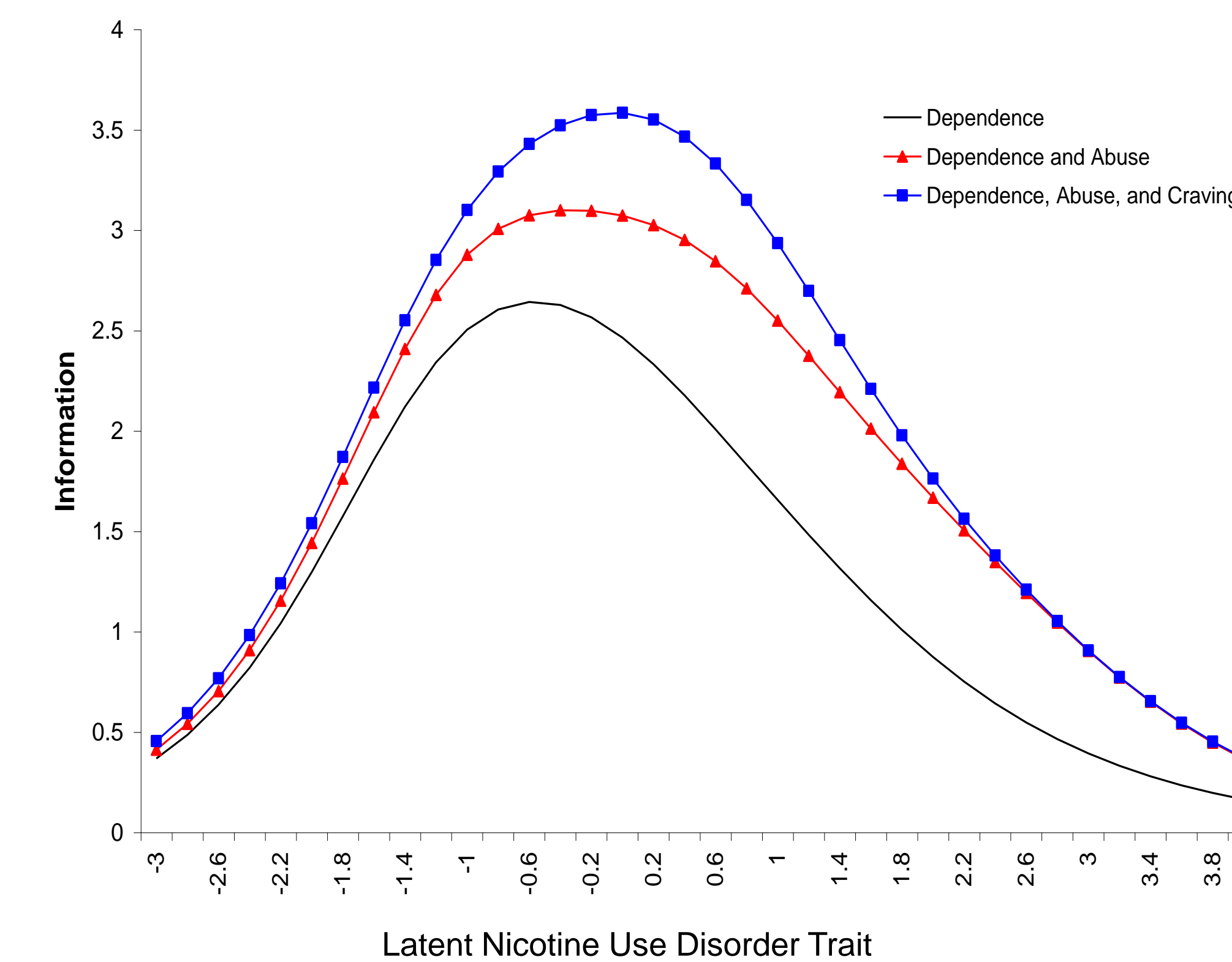
Criterion	%	In your ENTIRE LIFE did you EVER...
<b>Dependence</b>		
Tolerance	80.7	Find that you had to use much more tobacco than you once did to get the effect you wanted? OR Increase your smoking by at least 50 percent? OR Find that the first cigarette of the day had a much stronger effect than it used to? OR Find that you no longer got dizzy or nauseous from smoking?
Withdrawal	24.5	Withdrawal syndrome (4 or more symptoms that cause distress or dysfunction) OR smoke to avoid having any of these symptoms?
Larger/ longer	55.3	Have a period when you often smoked more than you intended to?
Quit/control	61.4	Want to stop or cut down on your smoking, regardless of whether or not you actually tried, more than once? OR Find that you were unable to stop or cut down on your smoking, more than once?
Time spent	38.5	Find yourself chain smoking?
Activities given up	16.0	Give up or cut down on activities that you were interested in or that gave you pleasure or that were important to you – like associating with friends or relatives or attending social activities because smoking was not permitted at the activity?
Physical/ psychological	79.0	Continue to smoke even though you knew it was causing you a health problem or making a health problem worse? OR Continue to smoke even though made you jittery, anxious, or depressed?
<b>Abuse</b>		
Neglect roles	07.7	Find that your smoking interfered with taking care of your work, school work, or work at home?
Hazardous use	36.3	Smoke in a situation that increased your chances of getting hurt – like smoking in bed or smoking around flammable chemicals?
Social/ interpersonal	41.3	Continue to smoke even though it made other people like family members angry or unhappy?
<b>Cravings</b>		
Unbearable/ strong desire	52.4	When you have run out of cigarettes, do you always or often find it almost unbearable until you can get them? OR Do you always or often get a strong desire to smoke when you haven't smoked for a while?

## Study Sample

Household residents were selected from the Israeli population register; 727 were lifetime cigarette smokers. Of these, 81.7% were male, 28.5% were immigrants from the Former Soviet Union, 19.3% were 21-29 years old, 33.6% were 30-44, and 47.2% were 45 years or older.

## IRT Total Information Curves

The total information curve (TIC) summarizes the combination of all the criterion and provides a visual inspection of **where** along the disorder continuum the criteria measure the disorder well (i.e. with more precision).



**Figure 3.** The total information curves are obtained by summing the respective criterion information curves shown in Figure 2. Notice that adding abuse criteria provides more information on the higher end of the disorder severity and cravings provides more information in the middle range.

## IRT Total Information Area Index

By integrating the total information curve we can obtain a single number summary called the Total Information Area (TIA) Index. A larger TIA index indicates the criteria set provides more information about the disorder.

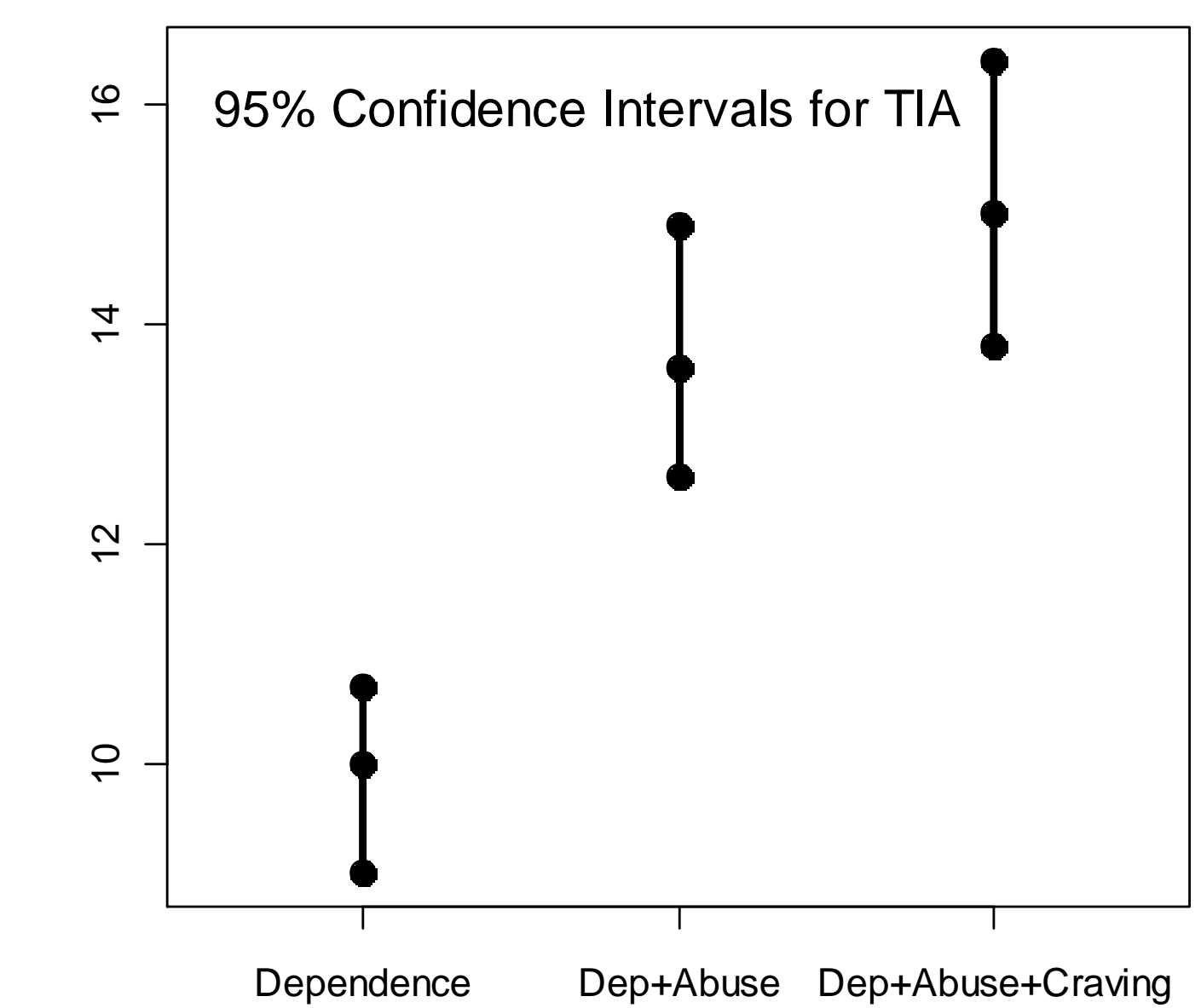
**Prior to our current work, there was no way to statistically test whether a larger TIA was indeed a statistically significant increase.**

The integral of the total information curve is equal to the sum of the individual criterion discrimination parameters. Thus, given maximum likelihood estimates of the discrimination parameters and their associated asymptotic covariance matrix (both available from IRT software), the delta method was used to create a closed-form estimate of the standard error of the TIA, allowing construction of 95% confidence intervals to compare across different criterion sets.

**Implementation was done in the R software.**

## Conclusions

Adding abuse criteria to the dependence criteria significantly increased the TIA and thus improves measurement of the disorder. The TIA is not further significantly improved by incorporating craving.



**Figure 4.** 95% confidence intervals for the TIA index for the 3 different criterion sets. **Larger values mean more information** in the criteria set for measuring the disorder. The confidence interval for Dep+Abuse is significantly higher (does not overlap) than the Dependence criteria alone.

When examining results from IRT models it is important to consider that results depend on sampling variability. We have provided a statistical method for properly taking into account that variability to help guide decisions between different criterion sets.

As DSM 5 moves towards continuous measures of disorders, we suggest keeping in mind the relevant parts of the continuum for clinical decision-making. A blind focus on the TIA index without considering the shape of the total information curve along the continuum could lead to choosing a criterion set with a better overall aggregate measurement that poorly measures the part of the continuum most relevant to clinical practice.

## Acknowledgments

This research was funded by R01AA013654, R01DA018652, K05AA014223 (Hasin), and the New York State Psychiatric Institute (Hasin). We would like to acknowledge the helpful consultations of Rachel Bar-Hamburger, Rina Meyer and Zalman Shoval in collecting the data in Israel. We also thank our collaborators: Dvora Shmulewitz, Katherine Keyes, Eliana Greenstein, Efrat Aharonovich, Christina Aivadyan, Baruch Spivak, Abraham Weizman, Amos Frisch and Ernest Poh

## For further information

About IRT Total Information:  
<http://www.columbia.edu/~mmw2177/irtprog.html>

About DSM-5 substance use disorders:  
<http://www.dsm5.org/Pages/Default.aspx>