

Validation of the Current Withdrawal Scale (CWS) to the Rasch Measurement Model,
GAIN Methods Report 1.0

Kendon J. Conrad
University of Illinois at Chicago

Karen M. Conrad
University of Illinois at Chicago

Barth B. Riley
Chestnut Health Systems

Rod Funk
Chestnut Health Systems
Michael L. Dennis
Chestnut Health Systems

Abstract

Purpose. The purpose of this report is to provide a brief psychometric analysis of the *Current Withdrawal Scale (CWS)* using the Rasch measurement model. The *CWS* consists of 21 past-week yes/no items related to psychological and physiological withdrawal symptoms based on the *Diagnostic and Statistical Manual of Mental Disorders-IV-TR* symptoms

Methods. Rasch analysis included an examination of: person and item reliabilities; construct validity including item and person fit statistics; and differential item functioning (DIF) across subgroups. DIF analysis allowed us to determine if the relative item estimates (i.e., item difficulty estimates) remained invariant across subgroups of persons. The person fit group analysis allowed us to examine the expected and unexpected patterns of raw scores in terms of the expectations of the Rasch model. The response patterns for each fit group are discussed relative to the response pattern for the overall mean. Data were collected on 18,317 persons who presented for substance problem screening; however, Rasch analysis is performed on individuals with non-extreme scores (i.e., the *Winsteps* software program deletes those with extreme minimum and maximum scores). Thus, data were analyzed on the 5353 persons with non-extreme scores, except in the case of the fit group analysis which was conducted on 5340 persons. The data set included a self-report variable asking respondents about the primary drug used. The response options examined included alcohol, amphetamines, cocaine, marijuana, opioids, and all other drugs.

Results. The *CWS* performs well as a measure of the construct of current withdrawal symptoms related to alcohol and drug use. Items appear to form a fairly reasonable hierarchy of symptom severity. Rasch person internal consistency reliability is .79 and item reliability is 1.00. The persons' responses generally conformed to the expectations of the Rasch model. Eight of the 21 *CWS* items exhibited significant DIF (i.e., $> .5$ SD = .70 logits). Significant DIF occurred in 3 items for youth vs. adults, 2 items for race when using African American as the reference group, and 7 items for primary substances when using opioids as the referent. There was no significant DIF between males and females. One item (*S3C11Hungrier*) misfit using the criterion of >1.33 MNSQ for both infit and outfit. In terms of model fit for persons, 83.4% of the sample showed a pattern of responses consistent with Rasch model expectations. Relative to the overall mean, those with atypical responses split into three groups: Atypical Groups 1, 2, and 3. The Atypical Type 1 group showed a deceptive pattern because their response patterns were lower than expected (i.e., lower than group average) on the less serious symptoms such as "feel tired"; "trouble sleeping" but higher than expected for the more serious symptoms, such as "see, feel, and hear things

that are not real"; *"have convulsions or seizures"*." Atypical Group Types 2 and 3 also tended to be deceptive because, even though they tended to have higher than expected scores, they tended to endorse the most severe symptoms more than expected but endorsed the lower severity items less than expected.

Conclusion. The CWS performs well as a measure of the construct of current withdrawal symptoms related to alcohol and drug use with good Rasch person internal consistency reliability and item reliability. The results suggest an insignificant impact of DIF for age and race because of the small number of items involved and the lack of theoretical implications for those items. There was no DIF for gender. However, further consideration should be given to the interpretation of the CWS for opioid users. Importantly, the fit group analysis showed that the CWS underestimates symptom severity in all three atypical groups. Relative to what was expected based on their CWS measures, people in the Atypical Type 1 group tended to score lower than expected mostly due to their tendency not to endorse low severity items. While people in the Atypical Types 2 and 3 groups tended not to endorse the low severity items, they tended to endorse the high severity items much more than expected. We recommend flagging the Atypical Groups 1, 2, and 3 for clinicians because these groups will tend to have scores that underestimate their current withdrawal.

Citation. Conrad, K. J., Conrad, K. M., Riley, B.B., Funk, R. & Dennis, M. L. (2010). Validation of the Current Withdrawal Scale (CWS) to the Rasch Measurement Model, GAIN Methods Report 1.0. Chicago, IL: Chestnut Health Systems. Retrieved from TBA

Purpose

The purpose of this report is to provide a brief report of a psychometric analysis of the *Current Withdrawal Scale (CWS)* using the Rasch measurement model (Bond & Fox, 2007; Rasch, 1960). The Rasch analysis was conducted using *Winsteps* software (Linacre, 2009). The tables and figures are annotated to summarize the main points.

Background

The *CWS* is a scale that is part of the larger *Global Appraisal of Individual Needs (GAIN)*. The *GAIN* is a biopsychosocial assessment that integrates research and clinical practice to do diagnosis, placement, individualized treatment planning, and program evaluation and meets major reporting requirements. The *GAIN* is a series of related instruments that share the same general instructions, questions (and variable names in most cases), scoring, interpretation, and clinical decision trees.

The *CWS* is a count of past-week symptoms related to psychological and physiological withdrawal symptoms associated with stopping, cutting down, or trying to limit one's use of alcohol or other drugs.

“Higher scores represent more physiological symptoms and suggest the increasing need for medical evaluation and monitoring. This scale was created by listing out the general and unique symptoms of withdrawal for each substance in DSM-IV, then working with a panel of clinicians and researchers to produce lay word descriptions. Symptoms are asked once, across substance.” (Chestnut Health Systems, [*GAIN*] v 5.6.2).

The Current Withdrawal Scale

The *CWS* consists of 21 items. The scale was asked to those respondents who answered in the affirmative to the following question, “During the past week, did you stop, try to stop, cut down or try to limit your use of alcohol or other drugs?” The item stem reads: “*When you did this, did you have any of the following **withdrawal symptoms or problems**? **Withdrawal symptoms** are a group of physical or emotional symptoms that happen when a person who regularly uses alcohol or other drugs stops using them, tries to stop using them, or cuts down on their use **during the past week.***”

The response format is Yes/No (coded: no=0, yes=1). The item stems, *GAIN* item numbers, Rasch output item location codes, and item labels are shown in the table below.

Table 1. CWS Scale and Item Information

Subscale Name and Item Stem	GAIN Item Number	Rasch Output Item Number	Item Label
<i>Current Withdrawal Scale</i>			
Move and talk much slower than usual	S3c1.	3.	S3C1MoveTalkSlower
Yawn more than usual	S3c2.	4.	S3C2Yawn
Feel tired	S3c3.	5.	S3C3FeelTired
Have bad dreams that seemed real	S3c4.	6.	S3C4BadDreamsSeemedReal
Have trouble sleeping, including sleeping too much or not being able to sleep	S3c5.	7.	S3C5TroubleSleeping
Feel sad, tense or angry	S3c6.	8.	S3C6FeelSadTenseAngry
Feel really nervous or tense	S3c7.	9.	S3C7FeelNervousTense
Fidget, pace, wring your hands or have trouble sitting still	S3c8.	10.	S3C8FFidgetPace
Have shaky hands	S3c9.	11.	S3C9ShakyHands
Have convulsions or seizures	S3c10.	12.	S3C10ConvulsionsSeizures
Feel hungrier than usual	S3c11.	13.	S3C11Hungrier
Throw up or feel like throwing up	S3c12.	14.	S3C12ThrowUp
Have diarrhea	S3c13.	15.	S3C13Diarrhea
Have muscle aches	S3c14.	16.	S3C14MuscleAches
Have a runny nose or eyes watering more than usual	S3c15.	17.	S3C15RunnyNoseEyesWater
Sweat more than usual, have your heart race or goose bumps	S3c16.	18.	S3C16SweatHeartRaceGooseBumps
Have a fever	S3c17.	19.	S3C17Fever
See, feel or hear things that are not real	S3c18.	20.	S3C18SeeFeelHeardThingsNotReal
Forget a lot of things or have problems remembering	S3c19.	21.	S3C19ForgetProblemsRemembering
Have any of these withdrawal problems kept you from doing social, family, job or other activities?	S3c20.	22.	S3C20WithdrawalKeptFromActivities
Have you used the same or another drug to stop or avoid having any of these withdrawal symptoms?	S3c21.	23.	S3C21UUseDrugStopWithdrawal

Data Source

Data on the 18,317 cases reported in this paper came from 12 projects/programs including 70 sites from around the country. All interviews were conducted by interviewers with three to four days of training followed by rigorous field-based certification procedures. Field interviewers had ongoing supervision by local trainers who were trained and certified by Chestnut staff on the use of the *GAIN*.

Full details about the *CWS* may be obtained at the following:

<http://www.chestnut.org/LI/gain/index.html>

Rasch Analysis

The Rasch measurement model (Rasch, 1960) was chosen for this analysis because it is the only item response theory model that has the desirable scaling properties of linear, interval measurement (Embretson & Reise, 2000). Therefore, Rasch measures are the most valid for mathematical operations, such as correlation and regression analysis, as well for assessing change. Rather than tailor models to fit the data, the Rasch one parameter model fulfills the requirements of fundamental measurement, i.e., linear interval scale (Bond & Fox, 2007), and examines the data, i.e., items and persons, for flaws or problems that are indicated by their failure to fit the model.

Quality control with fit statistics. Rasch analysis provides fit statistics to test assumptions of fundamental measurement (Wright & Stone, 1979). “Fitting the model” simply means meeting basic assumptions of measurement, e.g., high scorers should endorse or get right almost all of the easy items. Once identified, persons and items that “misfit” can then be examined qualitatively to determine the causes of the problems. Problems may include items with confusing wording or items that assess a construct that is different from the principal one being measured, i.e., multidimensionality. Understanding poor fit can lead to improving or dropping items.

The fit of the data to the model is evaluated by fit statistics that are calculated for both persons and items. The following link provides a handy guide to interpreting fit statistics: <http://www.rasch.org/rmt/rmt82a.htm>. The Rasch model provides two indicators of misfit: infit and outfit. The infit is sensitive to unexpected behavior affecting responses to items near the person ability level and the outfit is outlier sensitive. Mean square fit statistics are defined such that the model-specified uniform value of randomness is 1.0 (Wright & Stone, 1979). Person fit indicates the extent to which the person’s performance is consistent with the way the items are used by the other respondents. Item fit indicates the extent to which the use of a particular item is consistent with the way the sample respondents have responded to the other items. For this type of analysis, values between .75 and 1.33 logits (log odd units) are considered acceptable (Wilson, 2005). In addition to fit statistics, principal component analysis of residuals is used to examine whether a substantial factor exists in the residuals after the primary measurement dimension has been estimated (Linacre, 1998; Smith, 2002).

Construct Validation

In Rasch analysis the item hierarchy that is created by the item difficulty estimates provides an indication of construct validity (Smith, 2001). The items should form a ladder of low severity symptoms on the bottom to high severity symptoms on the top.

In summary, the advantages of Rasch analysis are that:

- Standard errors differ across scores of items and persons, e.g., improved estimation of error in extreme scores.
- Enables shorter measures that are more reliable, e.g., eliminate bad items, and via computerized adaptive testing.
- Facilitates analysis of construct validity
- Enables comparable scoring across different measures, i.e., item and test equating.
- Unbiased estimates of item difficulties can be obtained from non-representative samples.
- Interval scale properties are achieved. How? Probabilities, or log odds, are used.
- Analysis of response category usefulness is enhanced.
- Analysis of person and item characteristics is enhanced through fit statistics.
- Enables analysis of item bias, a.k.a., differential item functioning
- Facets beyond persons and items that affect the measures may be estimated

For references to articles that illustrate the applications noted above, we recommend Conrad & Smith (2004). For a complete treatment of Rasch analysis, we recommend Bond & Fox (2007) which includes a glossary of Rasch measurement terminology. Terminology may also be accessed online via *Rasch Measurement Transactions* located at <http://www.rasch.org/rmt/>. The tables below are output from Winsteps (Linacre, 2007) with annotated explanations and interpretations.

Background Characteristics of the Sample

As shown in the following table, the sample was predominately under 18 years of age (88%) and male (71%). Almost half were Caucasian (42%), a fifth were African American (20%), and the remainder Hispanic or mixed race. Of the top five most severe primary drugs reported, marijuana was reported by over 50% of the sample. The drug least often reported was opioids at 3%. Other drugs reported included amphetamines (16%), cocaine (6%), and alcohol (19%). Slightly over 4% percent of the sample reported other primary drugs.

Table 2. Demographic Characteristics of the Sample (N=18,317^a)

	Percent	Number
Age, Mean (sd) 17.63 (6.39)		
< 18 years	87.5	15,979
≥18 years	12.5	2,275
Gender		
Male	70.6	12,889
Female	29.4	5,358
Race		
African American	19.8	3,602
Caucasian	41.7	7,596
Hispanic	20.4	,711
Mixed/other	18.1	3,296
Drug, primary, most severe		
Alcohol	19.0	3,475
Amphetamines	15.6	2,850
Marijuana	52.4	9,591
Cocaine	5.9	1,075
Opiates	2.8	517
Other drug	4.3	794

^a Numbers may not add up to 18,317 due to missing values

Table 3. Person and Item Reliability

SUMMARY OF 5353 MEASURED (NON-EXTREME) PERSONS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	6.4	21.0	-1.38	.66	1.00	.1	1.00	.1
S.D.	4.7	.4	1.57	.19	.21	.8	.90	.7
MAX.	20.0	21.0	3.89	1.24	2.19	4.4	9.90	6.8
MIN.	1.0	4.0	-3.62	.50	.47	-2.7	.09	-1.5
REAL RMSE	.72	ADJ.SD 1.40	SEPARATION 1.95	PERSON RELIABILITY .79				
MODEL RMSE	.69	ADJ.SD 1.41	SEPARATION 2.04	PERSON RELIABILITY .81				
S.E. OF PERSON MEAN = .02								

MAXIMUM EXTREME SCORE: 6 PERSONS
MINIMUM EXTREME SCORE: 12754 PERSONS
LACKING RESPONSES: 43 PERSONS
DELETED: 161 PERSONS
VALID RESPONSES: 99.8%

SUMMARY OF 18113 MEASURED (EXTREME AND NON-EXTREME) PERSONS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	1.9	21.0	-3.85	1.50				
S.D.	3.9	.2	1.82	.55				
MAX.	21.0	21.0	5.30	2.18				
MIN.	.0	1.0	-4.89	.50				
REAL RMSE	1.60	ADJ.SD .87	SEPARATION .54	PERSON RELIABILITY .23				
MODEL RMSE	1.60	ADJ.SD .87	SEPARATION .54	PERSON RELIABILITY .23				
S.E. OF PERSON MEAN = .01								

PERSON RAW SCORE-TO-MEASURE CORRELATION = .97 (approximate due to missing data)
CRONBACH ALPHA (KR-20) PERSON RAW SCORE RELIABILITY = .93 (approximate due to missing data)

SUMMARY OF 21 MEASURED (NON-EXTREME) ITEMS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	1628.4	18097.6	.00	.04	1.00	-.8	1.06	-.3
S.D.	853.1	5.4	1.34	.01	.11	4.4	.31	4.3
MAX.	3205.0	18107.0	3.59	.10	1.33	9.9	2.14	9.9
MIN.	126.0	18087.0	-2.03	.03	.88	-7.4	.80	-6.8
REAL RMSE	.04	ADJ.SD 1.34	SEPARATION 30.24	ITEM RELIABILITY 1.00				
MODEL RMSE	.04	ADJ.SD 1.34	SEPARATION 30.73	ITEM RELIABILITY 1.00				
S.E. OF ITEM MEAN = .30								

DELETED: 24 ITEMS
UMEAN=.000 USCALE=1.000
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.96 (approximate due to missing data)
112239 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 90271.34 with 106866 d.f. p=1.0000

- Person reliability is acceptable at .79
- Items are placed reliably on the ruler.
- Cronbach alpha =.93
- Cronbach's alpha is higher because it estimates extreme scores as measured perfectly, i.e., with no error.
- A separation value of 1.95 gives approximately two separation levels, thus splitting the persons into about 3 groups on the Rasch ruler.
- Note that the Rasch analysis using the *Winsteps* software program was conducted on the 5353 persons with non-extreme scores.
- Of the 18, 317 persons in the complete data set, 214 cases were deleted by *Winsteps*, leaving 18,113 extreme and non-extreme persons.

- Good item reliability of 1.00.
- Item separation is high at 30.24 meaning items are placed reliably on the ruler.

Figure 1. Item Wright Map



- Hierarchy of symptom severity appears reasonable. With some exceptions, serious symptoms, like convulsions/seizures, hallucinations, continued use of drugs to avoid withdrawal symptoms, and activity limitations are in the upper half of item map while less serious symptoms, like fatigue, yawning, and hunger are in the bottom half of item map.
- *DSM_IV_TR* does not discuss a hierarchy of symptom severity.
- Note, some items list > symptom with symptoms (e.g., S3C16 sweat, heart race, goose bumps)
- Person mean is -1.38 showing persons exhibit low severity.
- Big floor effect >12,000 persons
- Floor effect is expected in substance use disorder screening
- Some redundancy of items
- A crosswalk of the alcohol and drug withdrawal symptoms listed for each primary drug in the *DSM_IV_TR* and the *Merck Manual* by the CWS items is shown in the Appendix. The response pattern for the primary drugs did not always correspond to the suggested pattern of symptoms (data not shown).

ITEM STATISTICS

Table 4. Principal Component Analysis of Standardized Residual Correlations for Items

STANDARDIZED RESIDUAL variance (in Eigenvalue units)

	=		-- Empirical --	Modeled
Total raw variance in observations	=	34.4	100.0%	100.0%
Raw variance explained by measures	=	13.4	39.0%	37.6%
Raw variance explained by persons	=	6.1	17.6%	16.9%
Raw Variance explained by items	=	7.4	21.5%	20.6%
Raw unexplained variance (total)	=	21.0	61.0%	100.0%
Unexplnd variance in 1st contrast	=	1.6	4.6%	7.5%
Unexplnd variance in 2nd contrast	=	1.4	4.0%	6.5%
Unexplnd variance in 3rd contrast	=	1.2	3.6%	5.9%
Unexplnd variance in 4th contrast	=	1.2	3.4%	5.6%
Unexplnd variance in 5th contrast	=	1.2	3.4%	5.5%

CON-	TRAST	LOADING	MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
1		.54	-2.03	.95	.91	A 5	S3C3FeelTired
1		.53	-1.05	.99	.96	B 4	S3C2Yawn
1		.25	-.59	1.33	1.68	C 13	S3C11Hungrier
1		.24	.17	.95	.89	D 3	S3C1MoveTalkSlower
1		.24	.02	.98	.95	E 17	S3C15RunnyNoseEyesWater
1		.12	.32	.92	.86	F 16	S3C14MuscleAches
1		.07	1.65	.95	1.00	G 15	S3C13Diarrhea
1		.06	1.97	.93	1.04	H 19	S3C17Fever
1		.03	-1.49	1.23	1.37	I 21	S3C19ForgetProblemsRemembering
1		.02	3.59	1.03	2.14	J 12	S3C10ConvulsionsSeizures
1		.01	.01	.90	.85	K 18	S3C16SweatHeartRaceGooseBumps
1		.01	1.38	1.10	1.16	j 20	S3C18SeeFeelHeardThingsNotReal
1		-.53	-.70	.88	.80	a 9	S3C7FeelNervousTense
1		-.49	-1.69	.92	.86	b 8	S3C6FeelSadTenseAngry
1		-.37	-1.01	.95	.90	c 10	S3C8FFidgetPace
1		-.24	.00	.94	.88	d 11	S3C9ShakyHands
1		-.19	-1.84	.94	.89	e 7	S3C5TroubleSleeping
1		-.15	.53	1.07	1.02	f 23	S3C21UUseDrugStopWithdrawal
1		-.11	.34	.98	.91	g 22	S3C20WithdrawalKeptFromActivit
1		-.06	-.10	1.05	1.13	h 6	S3C4BadDreamsSeemedReal
1		-.04	.52	.96	.97	i 14	S3C12ThrowUp

- To judge the strength of the measurement dimension, we used the following internal guidelines for variance explained by the measure: $\geq 40\%$ is considered a strong measurement dimension (Linacre, 2006), $\geq 30\%$ is considered a moderate measurement dimension, and $\geq 20\%$ is considered a minimal dimension. The 20% criterion is taken from Reckase (1979).
- The variance explained by the measure is 39%.
- The fact that only 7.5% of the variance is explained by the first factor of residuals supports unidimensionality
- Thus, the scale is considered unidimensional using these guidelines.

- Re: Misfit
- One item exhibited Infit and Outfit Misfit: *S3C11Hungrier*
- Two items exhibited Outfit Misfit:
 - *S3C19ForgetProblemsRemembering* (1.37)
 - *S3C10ConvulsionsSeizures* (2.14)

Table 5. Most Unexpected Item Responses in Terms of Measure

ITEM	MEASURE	PERSON
		11 111 1111 111111
		22223199522338183009755511 6552109988877665544442
		91316976732922629765184498432958824485011909264433
		70060256272306843660755190433179337238270775728436
		02428205927066302260380082006304925959662427652188
		high-----
10	S3C8FFidgetPa	-1.01 h 1.....
9	S3C7FeelNervo	-.70 a 1.....
13	S3C11Hungrier	-.59 B 0000.....
3	S3C1MoveTalkS	.17 g 1.....
16	S3C14MuscleAc	.32 d 1.....1.....
22	S3C20Withdraw	.34 J 1.....1.....
14	S3C12ThrowUp	.52 j 1.....1.....1.....1.....
20	S3C18SeeFeelH	1.38 D 1....1..1.....1.1.111...1.11..
15	S3C13Diarrhea	1.65 H 1...1.....1111.....1....
19	S3C17Fever	1.97 G 111.1.11..1.1.....1.1...11...1.
12	S3C10Convulsi	3.59 A ...1111111111...1.1..11.1.1.....1.....1
		-----low-
		11223199521118111119755511431111119988877665544442
		22316976732332683005184498436552104485011909264433
		91060256272926829760755190002958827238270775728436
		70428205922306343660380082 3179335959662427652188
		02 706 0226 630492

This table shows the items and their measures on the left and the persons with the most unexpected responses on the right.

For example
We see that Person #6974 had a low measure of -3.62. In spite of this low measure, this person endorsed the high severity item *Convulsions* (i.e., 3.59).

Table 6. Most Unexpected Item Responses in Terms of OUTMSQ

ITEM	OUTMNSQ	PERSON
		11 111 1111 111111
		22223199522338183009755511 6552109988877665544442
		91316976732922629765184498432958824485011909264433
		70060256272306843660755190433179337238270775728436
		02428205927066302260380082006304925959662427652188
		high-----
12 S3C10Convulsi	2.14 A	...111111111111...1.1..11.1.1.....1.....1
13 S3C11Hungrier	1.68 B	0000.....
20 S3C18SeeFeelH	1.16 D1.....1..1.....1.1.111...1.11..
19 S3C17Fever	1.04 G111.1.11..1.1.....1.1....11....1.
15 S3C13Diarrhea	1.00 H1...1.....11111.....1....
22 S3C20Withdraw	.91 J1.....1.....
14 S3C12ThrowUp	.97 j1.....1...1...1.....
10 S3C8FFidgetPa	.90 h1.....
3 S3C1MoveTalkS	.89 g1.....
16 S3C14MuscleAc	.86 d1.....1.....
9 S3C7FeelNervo	.80 a1.....
		-----low-
		11223199521118111119755511431111119988877665544442
		22316976732332683005184498436552104485011909264433
		91060256272926829760755190002958827238270775728436
		70428205922306343660380082 3179335959662427652188
		02 706 0226 63049

This table shows the most misfitting items in terms of OUTMNSQ. To the right of the items, we see the persons who unexpectedly endorsed or did not endorse an item given their overall measure. *Item #12 S3C10Convulsions* has the highest outfit (2.14) because a few people with low overall measures endorsed it (the 1's) even though it was a high severity item. Also of note is *Item #13 S3C11Hungrier* where a few other people misfit because they had high overall measures, but did not endorse this item (the 0's - for example, Person 12970 has a high overall measure of 3.89 [data not shown]).

PERSON STATISTICS

Table 7. Misfit Order (table abridged due to space limitations)

ENTRY NUMBER	TOTAL		MEASURE	MODEL		INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		PERSON
	SCORE	COUNT		S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%		
440	2	21	-2.82	.78	1.18	.5	9.90	5.6	A-.25	.26	90.5	90.5	3273 1 2 0 2 2 2 1 1 1 3	
1683	3	21	-2.30	.67	1.46	1.1	9.90	5.0	B-.40	.31	85.7	85.7	2774 2 3 0 1 2 2 1 1 1 3	
1998	2	21	-2.82	.78	1.06	.3	9.90	5.6	C-.19	.26	90.5	90.5	14993 1 4 0 1 2 2 1 1 1 3	
2368	1	21	-3.62	1.05	1.21	.5	9.90	6.8	D-.60	.19	95.2	95.2	16598 2 4 0 1 2 2 1 1 1 3	
4338	1	21	-3.62	1.05	1.21	.5	9.90	3.3	E-.33	.19	95.2	95.2	6542 1 2 0 2 2 2 1 1 1 3	
5410	2	21	-2.82	.78	1.06	.3	9.90	5.6	F-.19	.26	90.5	90.5	18010 1 1 0 1 1 2 1 1 1 3	
5957	1	21	-3.62	1.05	1.21	.5	9.90	3.3	G-.33	.19	95.2	95.2	8389 1 2 0 2 2 1 1 1 1 1	
6072	1	21	-3.62	1.05	1.21	.5	9.90	3.3	H-.33	.19	95.2	95.2	16172 1 4 0 1 2 2 1 1 1 3	
6974	1	21	-3.62	1.05	1.21	.5	9.90	6.8	I-.60	.19	95.2	95.2	14172 1 4 0 1 2 2 2 1 1 6	
7173	2	21	-2.82	.78	1.18	.5	9.90	5.6	J-.25	.26	90.5	90.5	922 1 2 0 2 2 2 1 1 0 3	
8266	3	21	-2.30	.67	1.17	.5	9.90	4.7	K-.09	.31	85.7	85.7	5928 2 2 0 2 2 2 1 1 0 3	
8589	1	21	-3.62	1.05	1.21	.5	9.90	3.3	L-.33	.19	95.2	95.2	14870 1 2 0 2 2 2 1 1 0 3	
9429	1	21	-3.62	1.05	1.21	.5	9.90	3.3	M-.33	.19	95.2	95.2	10285 2 4 0 1 2 2 1 1 0 3	
10666	2	21	-2.82	.78	1.39	.8	9.90	5.7	N-.48	.26	90.5	90.5	3671 1 3 0 1 2 2 1 1 1 3	
13206	3	21	-2.30	.67	1.09	.4	9.90	4.7	O-.05	.31	85.7	85.7	6558 2 1 1 1 1 2 2 1 1 4	
13930	3	21	-2.30	.67	1.46	1.1	9.90	5.0	P-.40	.31	85.7	85.7	2766 2 4 0 1 2 2 1 1 1 3	
16236	1	21	-3.62	1.05	1.21	.5	9.90	6.8	Q-.60	.19	95.2	95.2	14963 1 2 0 2 2 2 1 1 1 3	
4625	1	21	-3.62	1.05	1.20	.5	9.28	2.8	R-.28	.19	95.2	95.2	6056 1 1 0 1 1 2 1 1 0 3	
10232	1	21	-3.62	1.05	1.20	.5	9.28	2.8	S-.28	.19	95.2	95.2	5889 1 3 0 1 2 2 1 1 0 3	
11839	1	21	-3.62	1.05	1.20	.5	9.28	2.8	T-.28	.19	95.2	95.2	9326 1 2 0 2 2 2 1 1 1 3	
12894	1	21	-3.62	1.05	1.20	.5	9.28	2.8	U-.28	.19	95.2	95.2	7770 2 2 1 2 2 1 1 1 1 1	
15570	1	21	-3.62	1.05	1.20	.5	9.28	2.8	V-.28	.19	95.2	95.2	5294 1 2 0 2 2 2 1 1 0 3	
15913	1	21	-3.62	1.05	1.20	.5	9.28	2.8	W-.28	.19	95.2	95.2	10164 1 2 0 2 2 2 1 1 1 3	
12227	5	21	-1.55	.57	1.05	.3	8.79	4.3	X .12	.39	81.0	78.3	13854 1 1 1 1 1 2 2 1 1 4	
2372	6	21	-1.25	.54	1.88	2.9	8.06	4.7	Y-.39	.41	52.4	76.2	17234 1 4 0 1 2 2 2 1 1 2	
16745	3	12	-.74	.73	1.76	1.9	7.81	3.7	Z-.49	.38	66.7	78.0	2414 1 3 1 1 2 2 2 1 1 4	
1802	2	21	-2.82	.78	1.39	.8	7.18	2.7	- .30	.26	90.5	90.5	17928 1 4 0 1 2 2 1 1 1 3	
5858	2	21	-2.82	.78	1.39	.8	7.18	2.7	- .30	.26	90.5	90.5	853 1 2 0 2 2 1 1 1 0 1	
13932	2	21	-2.82	.78	1.39	.8	7.18	2.7	- .30	.26	90.5	90.5	3154 1 4 0 1 2 2 1 1 1 3	
4441	1	21	-3.62	1.05	1.20	.5	7.11	2.4	- .23	.19	95.2	95.2	701 1 1 0 1 1 2 1 1 1 3	
4482	1	21	-3.62	1.05	1.20	.5	7.11	2.4	- .23	.19	95.2	95.2	6527 1 2 0 2 2 2 1 1 1 3	
5276	1	21	-3.62	1.05	1.20	.5	7.11	2.4	- .23	.19	95.2	95.2	538 1 1 0 1 1 2 2 1 1 2	
7102	1	21	-3.62	1.05	1.20	.5	7.11	2.4	- .23	.19	95.2	95.2	9001 2 2 0 2 2 1 1 1 0 1	
7176	1	21	-3.62	1.05	1.20	.5	7.11	2.4	- .23	.19	95.2	95.2	1003 1 2 0 2 2 2 1 1 0 3	

Over 83% of the sample/data fit the Rasch model. See **Fit Group Analysis** section of the report for details

Table 8. Persons with Most Unexpected Responses in Terms of Measures

PERSON	MEASURE	ITEM
		2 1 1 111 12122111
		578140936187362430592
	high-----	
2304	11063	2 2 0 2 2 2 2 1 1 6
12102	8193	1 2 1 2 2 2 2 5 1 5
12970	14069	2 2 1 2 2 2 2 5 1 5
2162	8898	1 3 0 1 2 2 2 5 1 5
1922	9160	1 4 0 1 2 1 1 1 1 1
3608	3095	1 4 0 1 2 1 1 1 1 1
2372	17234	1 4 0 1 2 2 2 1 1 2
5729	17048	2 1 0 1 1 2 1 1 0 3
9665	17744	1 3 0 1 2 2 2 1 0 4
9750	12348	1 2 0 2 2 2 1 1 1 3
12227	13854	1 1 1 1 2 2 1 1 4
1683	2774	2 3 0 1 2 2 1 1 1 3
8266	5928	2 2 0 2 2 2 1 1 0 3
13206	6558	2 1 1 1 2 2 1 1 4
13930	2766	2 4 0 1 2 2 1 1 1 3
330	90	1 3 0 1 2 1 1 1 0 1
440	3273	1 2 0 2 2 2 1 1 1 3
1802	17928	1 4 0 1 2 2 1 1 1 3
1998	14993	1 4 0 1 2 2 1 1 1 3
5410	18010	1 1 0 1 1 2 1 1 1 3
5450	4081	1 4 0 1 2 2 1 1 0 3
5858	853	1 2 0 2 2 1 1 1 0 1
7173	922	1 2 0 2 2 2 1 1 0 3
9500	16135	1 3 0 1 2 2 1 1 0 3
10666	3671	1 3 0 1 2 2 1 1 1 3
10762	13965	1 3 0 1 2 2 1 1 1 3
13932	3154	1 4 0 1 2 2 1 1 1 3
18240	16511	2 4 0 1 2 2 1 1 1 3
2368	16598	2 4 0 1 2 2 1 1 1 3
4338	6542	1 2 0 2 2 2 1 1 1 3
4441	701	1 1 0 1 1 2 1 1 1 3
4482	6527	1 2 0 2 2 2 1 1 1 3
4625	6056	1 1 0 1 1 2 1 1 0 3
5276	538	1 1 0 1 1 2 2 1 1 2
5957	8389	1 2 0 2 2 1 1 1 1 1
6072	16172	1 4 0 1 2 2 1 1 1 3
6974	14172	1 4 0 1 2 2 2 1 1 6
7102	9001	2 2 0 2 2 1 1 1 0 1
7176	1003	1 2 0 2 2 2 1 1 0 3
8026	3325	2 2 0 2 2 2 1 1 1 3
8589	14870	1 2 0 2 2 2 1 1 0 3
8835	14817	2 4 0 1 2 2 2 1 0 6
9429	10285	2 4 0 1 2 2 1 1 0 3
9475	13293	1 3 0 1 2 2 1 1 0 3
10232	5889	1 3 0 1 2 2 1 1 0 3
11839	9326	1 2 0 2 2 2 1 1 1 3
12894	7770	2 2 1 2 2 1 1 1 1 1
15570	5294	1 2 0 2 2 2 1 1 0 3
15913	10164	1 2 0 2 2 2 1 1 1 3
16236	14963	1 2 0 2 2 2 1 1 1 3
		-----low-
		578241916111312122111
		1 0 3 187 62430592

- We see some of the high scorers with unexpected non-endorsement of low severity items, such as item #13-S3C11Hungrier (i.e., zeros on the left).
- The other unexpected responses are the lower scorers who are endorsing high severity items, especially #12-S3C10ConvulsionsSeizures

Table 9. Persons with Most Unexpected Responses in Terms of OUTFIT MSQ

PERSON	OUTMNSQ	ITEM
		2 1 1 111 12122111
		578140936187362430592
	high-----	
440 3273	1 2 0 2 2 2 1 1 1 3	9.90 A1
1683 2774	2 3 0 1 2 2 1 1 1 3	9.90 B1..1
1998 14993	1 4 0 1 2 2 1 1 1 3	9.90 C1
2368 16598	2 4 0 1 2 2 1 1 1 3	9.90 D1
4338 6542	1 2 0 2 2 2 1 1 1 3	9.90 E1.
5410 18010	1 1 0 1 1 2 1 1 1 3	9.90 F1
5957 8389	1 2 0 2 2 1 1 1 1 1	9.90 G1.
6072 16172	1 4 0 1 2 2 1 1 1 3	9.90 H1.
6974 14172	1 4 0 1 2 2 2 1 1 6	9.90 I1
7173 922	1 2 0 2 2 2 1 1 0 3	9.90 J1
8266 5928	2 2 0 2 2 2 1 1 0 3	9.90 K1
8589 14870	1 2 0 2 2 2 1 1 0 3	9.90 L1.
9429 10285	2 4 0 1 2 2 1 1 0 3	9.90 M1.
10666 3671	1 3 0 1 2 2 1 1 1 3	9.90 N1.....1
13206 6558	2 1 1 1 1 2 2 1 1 4	9.90 O1
13930 2766	2 4 0 1 2 2 1 1 1 3	9.90 P1..1
16236 14963	1 2 0 2 2 2 1 1 1 3	9.90 Q1
4625 6056	1 1 0 1 1 2 1 1 0 3	9.28 R1..
10232 5889	1 3 0 1 2 2 1 1 0 3	9.28 S1..
11839 9326	1 2 0 2 2 2 1 1 1 3	9.28 T1..
12894 7770	2 2 1 2 2 1 1 1 1 1	9.28 U1..
15570 5294	1 2 0 2 2 2 1 1 0 3	9.28 V1..
15913 10164	1 2 0 2 2 2 1 1 1 3	9.28 W1..
12227 13854	1 1 1 1 1 2 2 1 1 4	8.79 X1.....1
2372 17234	1 4 0 1 2 2 2 1 1 2	8.06 Y1.1..1.1
16745 2414	1 3 1 1 2 2 2 1 1 4	7.81 Z1.1
	-----low-	
		578241916111312122111
		1 0 3 187 62430592

- This table shows the persons who unexpectedly endorsed an item given their overall measure. Most of the person misfit is being caused by Items 12, 15, 19 when persons with low severity are unexpectedly endorsing the high severity items.
- Person #16236 is one of the most misfitting persons with a low measure (-3.62) but an unexpected endorsement of item #12.

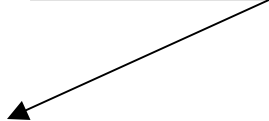
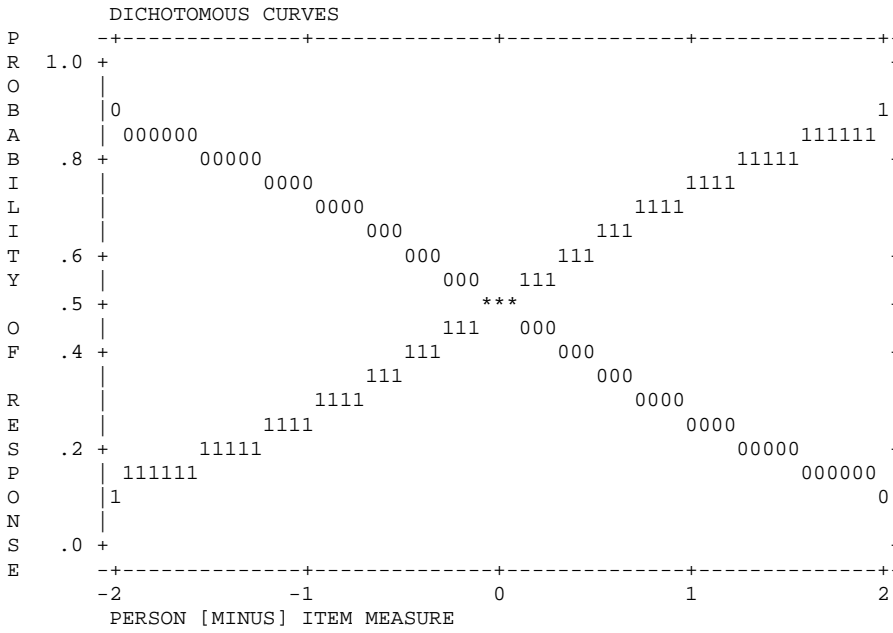


Table 10. Summary of Category Structure

CATEGORY	OBSERVED	OBSVD	SAMPLE	INFINIT	OUTFIT	COHERENCE	ESTIM		
LABEL	SCORE	COUNT	%	AVRGE	EXPECT	MNSQ	MNSQ	M->C C->M	DISCR
0	0	78109	70	-2.18	-2.18	1.00	1.01	84% 90%	0
1	1	34130	30	.44	.44	1.00	1.08	73% 61%	1.00
MISSING		174	0	-.77					

70% of the responses were zeros (did not endorse items)



Differential Item Functioning for Age, Gender, Race, and Primary Drug Severity for the Current Withdrawal Scale

As Bond and Fox (2007) note, the Rasch model requires that relative item estimates (i.e., item difficulty estimates) remain invariant across subgroups of persons (e.g., females and males). DIF allows us to examine whether items have significantly different meanings for different groups. They suggest that items that show DIF should be investigated to determine what may be inferred about the underlying construct and what that implies about the samples of persons detected. A significant DIF contrast is based on $> 1.4/2 = .7$ logit difference for all comparisons which is approximately half a standard deviation (Table 3, top panel) for the persons (Norman, Sloan, & Wyrwich, 2003; Conrad, Dennis, Bezruczko, Funk, & Riley, 2007).

The figures below present easily interpretable graphs of the relationships of the various groups on the items. Table 12 contains the data that formed these graphs, and provides the information to compute the exact differences between groups on each item. For example, to get the DIF contrast between males and females on *MoveTalkSlower*, you would subtract female = .33 from male = .08 to get a DIF contrast = -0.25, which is a non-significant difference.

Gender DIF. In Figure 2, we can see that there were no significant differences between males and females.

Age DIF. In Figure 3, there were significant differences on three items. *S3C19ForgetProblemsRemembering* and *S3C18SeeFeelHeardThingsNotReal* were endorsed more by youth. *S3C13Diarrhea* was endorsed more by adults.

Race DIF. In looking carefully at Figure 4, we observed that African-Americans tended to differ most from the other groups and were usually most different from Caucasians. It was more difficult for African-Americans to endorse *S3C9ShakyHands*. It was easier for African-Americans to endorse *S3C13Diarrhea*.

Primary Drug Severity DIF. As shown in Figure 5, with six different types of drugs, we observed that opioid users tended to differ most from the other groups in general, but they did not tend to be the most different from any particular group of users. It was easier for opioid users to endorse *S3C13Diarrhea*, *S3C14MuscleAches*, *S3C15RunnyNoseEyesWater*, and *S3C16SweatHeartRaceGooseBumps*. It was more difficult for opioid users to endorse *S3C11Hungrier*, *S3C18SeeFeelHeardThingsNotReal*, and *S3C19ForgetProblemsRemembering*. This pattern of symptoms for opioids is consistent with current clinical knowledge. (See the Appendix which contains a table that is a crosswalk of *DSM-IV_TR* and the *Merck Manual* drug withdrawal symptoms by the CWS).

Figure 2. CWS Gender DIF

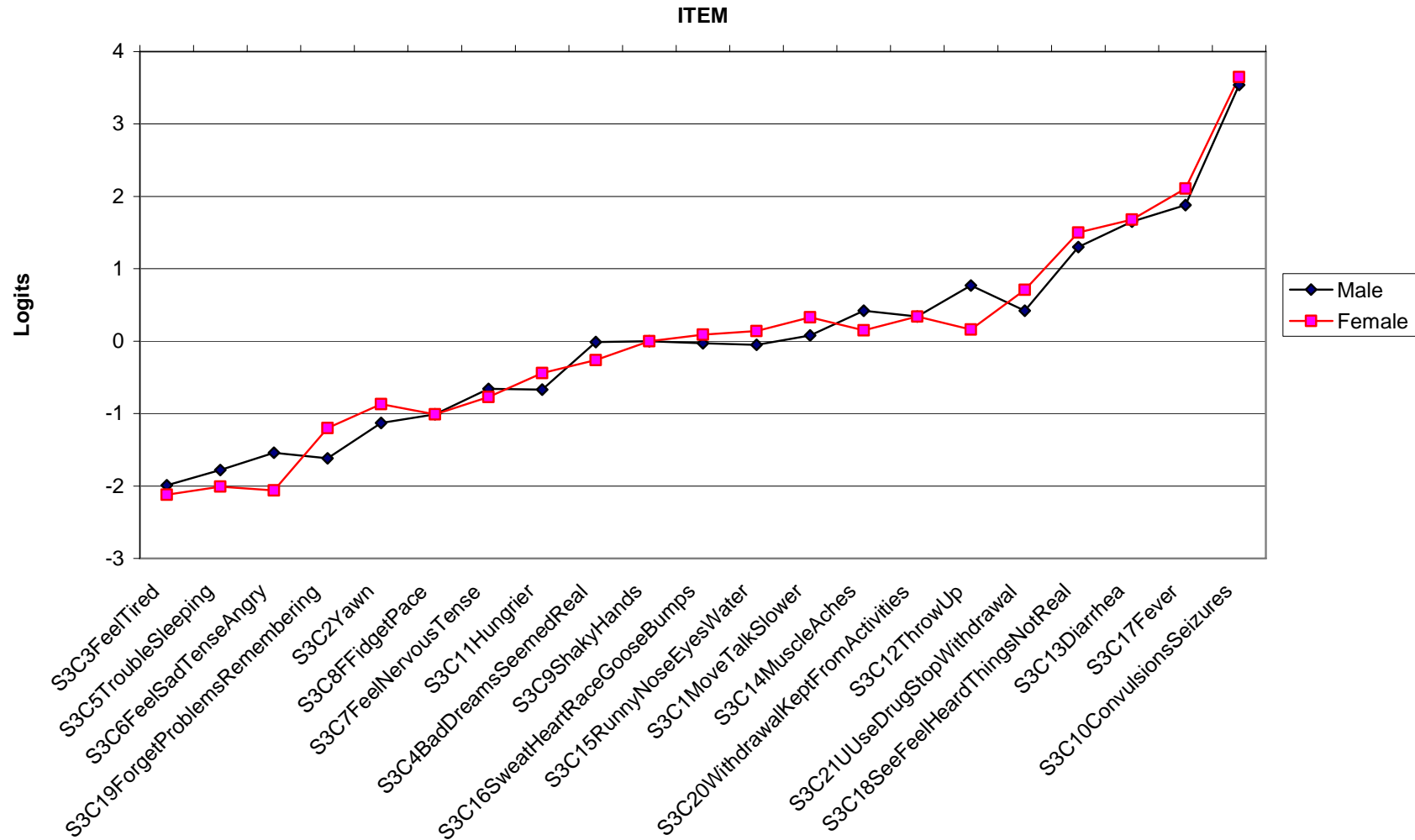


Figure 3. CWS Age DIF

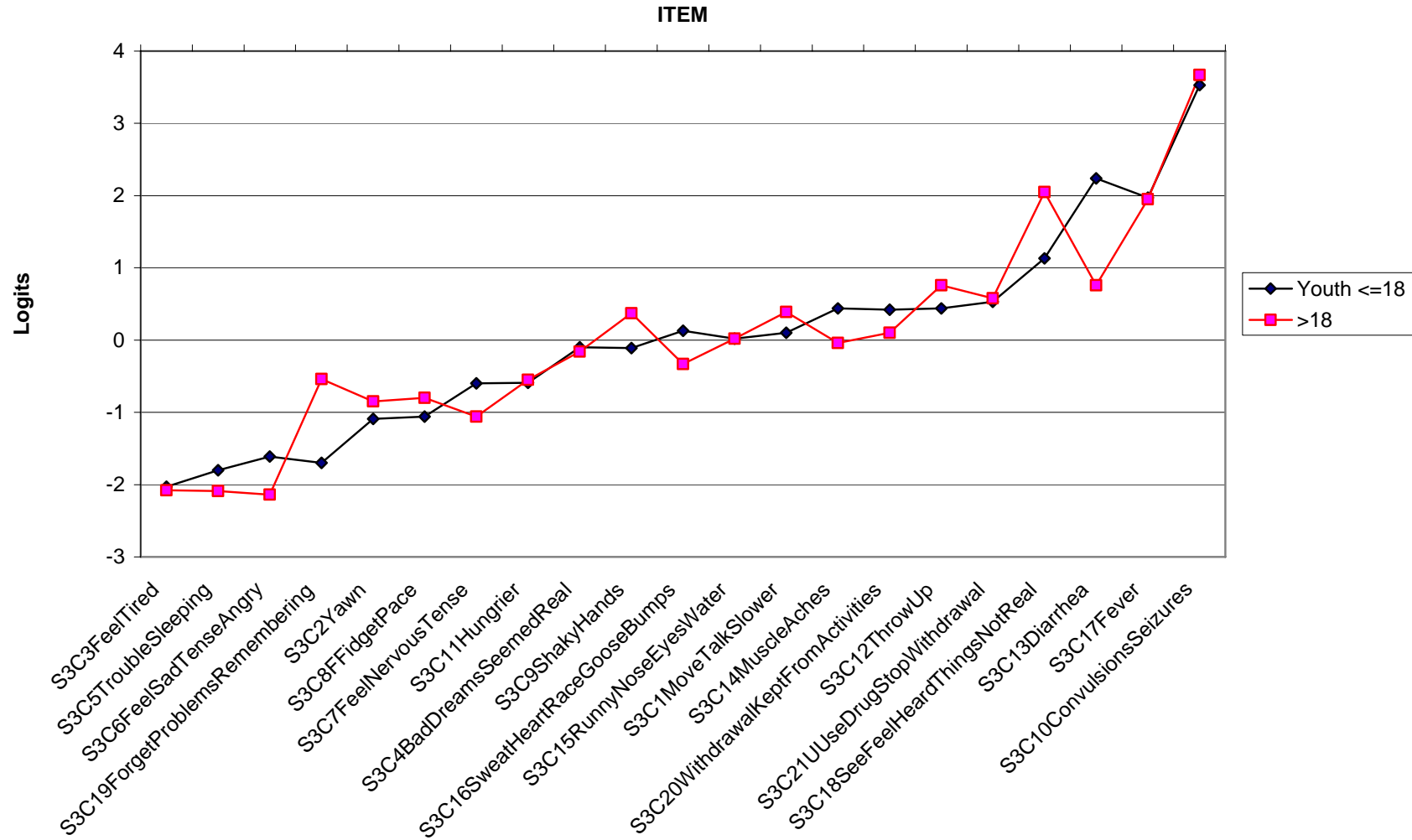


Figure 4. CWS Race DIF

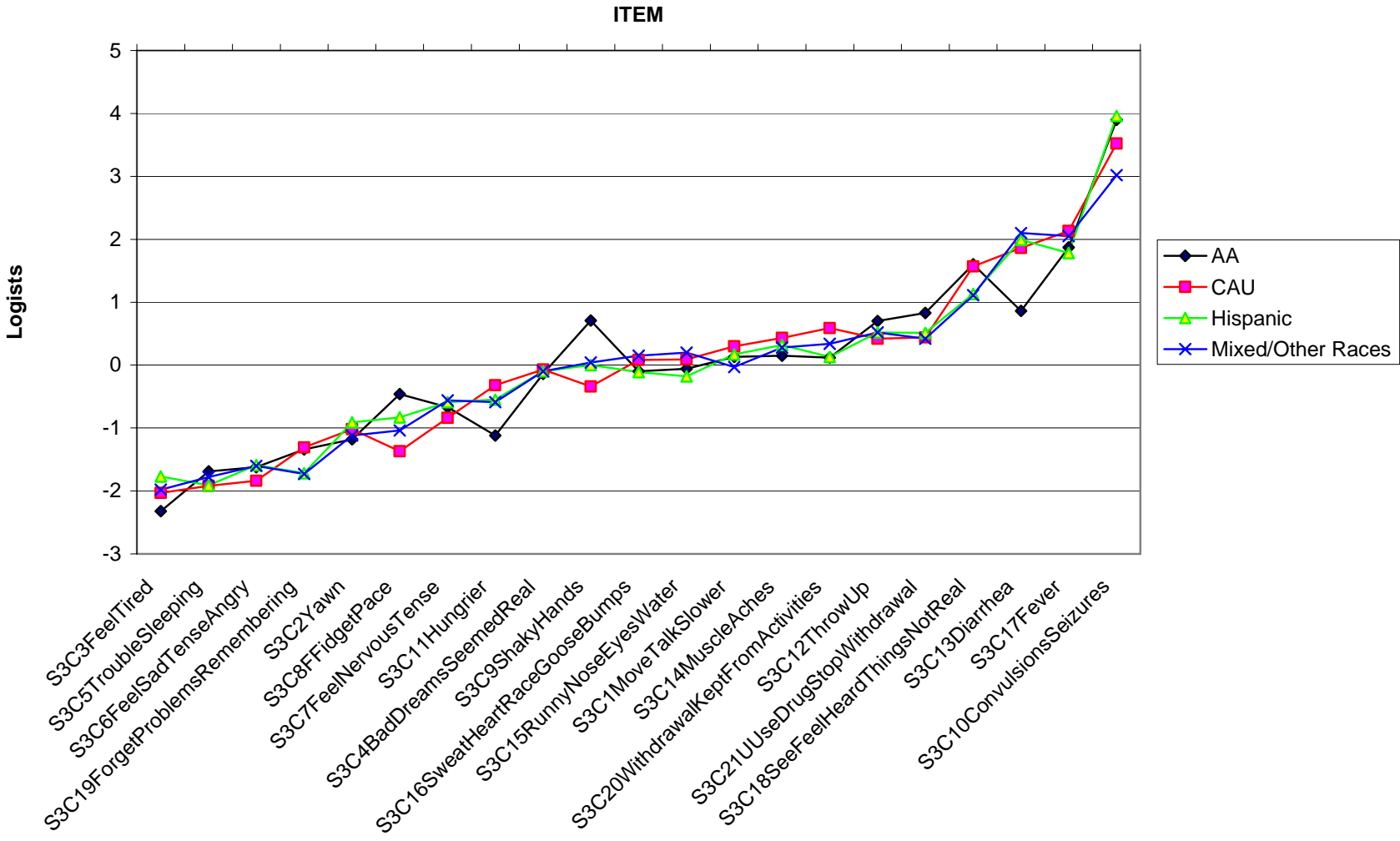


Figure 5. CWS Primary Drug DIF

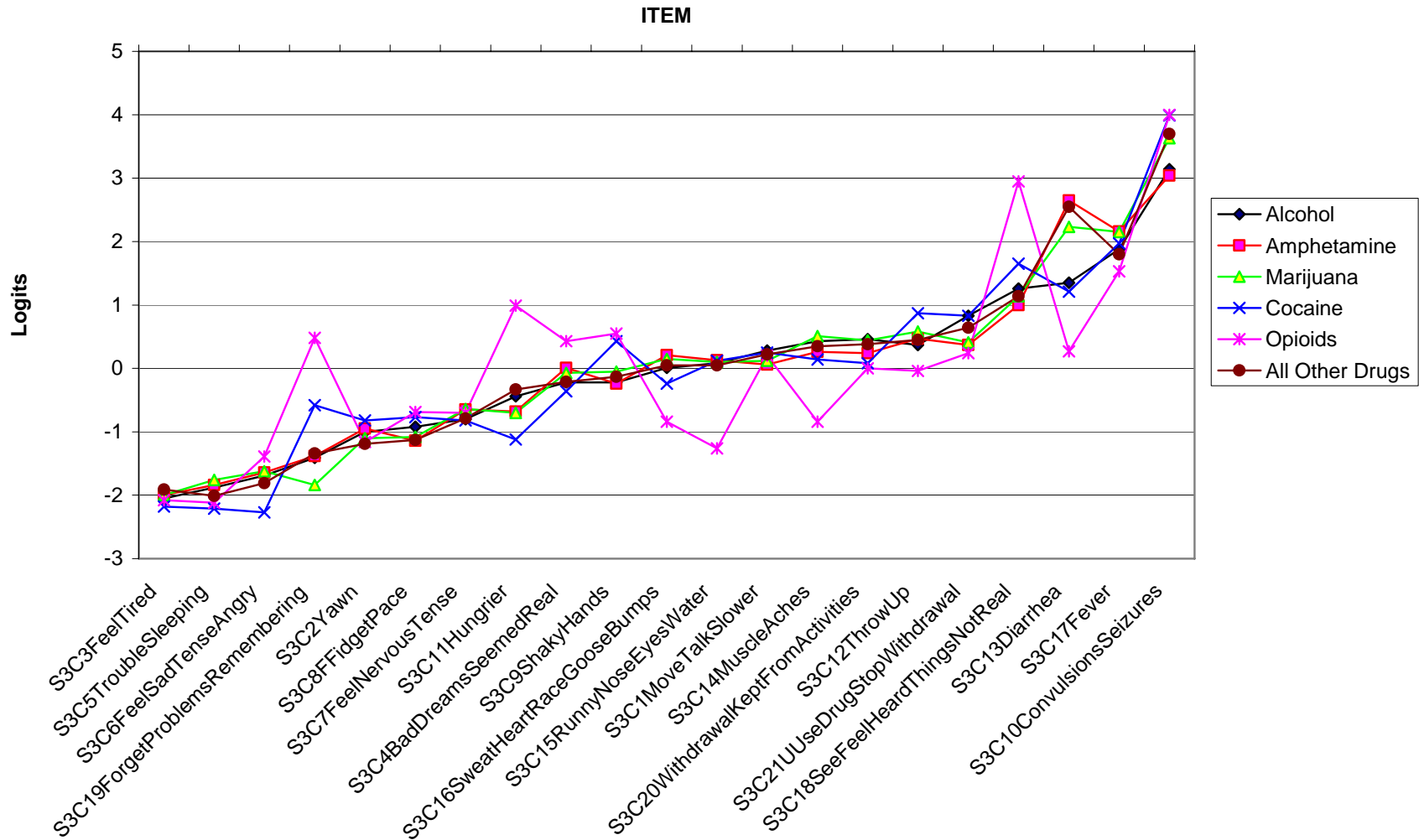


Table 11. CWS Item Measures by Demographic Groups (Items Listed in Severity Order from Easy to Hard to Endorse)

CWS ITEM NAME	CWS RASCH MEASURE	Gender		Age		Race				Primary Drug					
		MALE	FEMALE	YOUTH <18	ADULT >18	AA	CAU	HISP	MIXED/ OTHER	ALC	AMP	MAR	COC	OPIOIDS	OTHER DRUGS
S3C3FeelTired	-2.03	-1.99	-2.12	-2.03	-2.08	-2.32	-2.03	-1.77	-1.98	-2.05	-2	-2	-2.18	-2.08	-1.91
S3C5TroubleSleeping	-1.84	-1.78	-2.01	-1.8	-2.09	-1.69	-1.92	-1.91	-1.78	-1.88	-1.84	-1.76	-2.21	-2.12	-2.01
S3C6FeelSadTenseAngry	-1.69	-1.54	-2.06	-1.61	-2.14	-1.62	-1.84	-1.59	-1.6	-1.69	-1.64	-1.62	-2.27	-1.39	-1.81
S3C19ForgetProblemsRemembering	-1.49	-1.62	-1.2	-1.7	-0.54	-1.34	-1.31	-1.72	-1.73	-1.41	-1.38	-1.84	-0.58	0.48	-1.34
S3C2Yawn	-1.05	-1.13	-0.87	-1.09	-0.85	-1.18	-1.02	-0.91	-1.12	-1	-0.95	-1.1	-0.82	-1.16	-1.19
S3C8FFidgetPace	-1.01	-1.01	-1.01	-1.06	-0.8	-0.46	-1.37	-0.83	-1.04	-0.92	-1.14	-1.08	-0.77	-0.69	-1.13
S3C7FeelNervousTense	-0.7	-0.66	-0.77	-0.6	-1.06	-0.67	-0.84	-0.59	-0.56	-0.8	-0.65	-0.64	-0.82	-0.7	-0.79
S3C11Hungrier	-0.59	-0.67	-0.44	-0.59	-0.55	-1.12	-0.32	-0.55	-0.59	-0.44	-0.68	-0.7	-1.12	0.99	-0.33
S3C4BadDreamsSeemedReal	-0.1	-0.01	-0.26	-0.1	-0.16	-0.14	-0.07	-0.1	-0.1	-0.22	0.01	-0.07	-0.36	0.43	-0.21
S3C9ShakyHands	0	0	0	-0.11	0.37	0.71	-0.34	0	0.04	-0.22	-0.24	-0.05	0.43	0.55	-0.13
S3C16SweatHeartRaceGooseBumps	0.01	-0.03	0.09	0.13	-0.33	-0.1	0.08	-0.11	0.15	0.01	0.21	0.15	-0.24	-0.84	0.05
S3C15RunnyNoseEyesWater	0.02	-0.05	0.14	0.02	0.02	-0.06	0.09	-0.18	0.2	0.08	0.13	0.1	0.12	-1.26	0.05
S3C1MoveTalkSlower	0.17	0.08	0.33	0.1	0.39	0.13	0.3	0.17	-0.03	0.28	0.06	0.12	0.25	0.2	0.22
S3C14MuscleAches	0.32	0.42	0.15	0.44	-0.04	0.15	0.43	0.32	0.28	0.43	0.26	0.51	0.14	-0.84	0.35
S3C20WithdrawalKeptFromActivities	0.34	0.34	0.34	0.42	0.1	0.12	0.59	0.13	0.34	0.46	0.24	0.44	0.08	0	0.38
S3C12ThrowUp	0.52	0.77	0.16	0.44	0.76	0.7	0.42	0.52	0.52	0.37	0.47	0.58	0.87	-0.04	0.45
S3C21UUseDrugStopWithdrawal	0.53	0.42	0.71	0.53	0.58	0.83	0.44	0.51	0.42	0.83	0.37	0.41	0.83	0.24	0.64
S3C18SeeFeelHeardThingsNotReal	1.38	1.3	1.5	1.13	2.05	1.6	1.57	1.13	1.11	1.26	1	1.14	1.65	2.95	1.14
S3C13Diarrhea	1.65	1.65	1.68	2.24	0.76	0.86	1.86	1.99	2.1	1.35	2.65	2.23	1.21	0.27	2.55
S3C17Fever	1.97	1.88	2.11	1.97	1.95	1.87	2.13	1.78	2.05	1.88	2.16	2.15	1.97	1.53	1.8
S3C10ConvulsionsSeizures	3.59	3.54	3.65	3.53	3.67	3.9	3.52	3.96	3.02	3.14	3.04	3.63	3.99	4	3.7

Person Fit Group Analysis for the CWS

The purpose of the person fit group analysis is to illustrate and interpret the expected and unexpected patterns of raw scores in terms of the expectations of the Rasch model. The response patterns for each fit group are discussed relative to the response pattern for the overall mean.

This information should enable us to interpret certain unusual patterns of scores more appropriately, e.g., low scorers who do not have less severe withdrawal symptoms, such as *FeelTired*, *TroubleSleeping*, *FeelSadTenseAngry*, and *ForgetProblemsRemembering*, but who are actually at high risk because of their severe withdrawal symptoms, such as *UseDrugStopWithdrawal*, *SeeFeelHeardThingsNotReal*, *Diarrhea*, *Fever*, and *ConvulsionsSeizures*. This type of analysis should inform the interpretation of Rasch measures and enable better treatment decisions. We alert the reader that these charts present raw score p-values (higher proportions endorsing the items are higher on the charts) so they are upside down from the typical Rasch charts where more rare is higher. We use the words **lower, moderate and higher to express** where the fit group means fall relative to the overall group mean. The items in the charts are arranged in severity order.

Figure 6 displays the numbers and percentages of persons in each fit group in a pie chart.

In Figures 7-10 below, we present the four possible person fit patterns using Rasch person fit statistics (Wright & Stone, 1979) where ≤ 1.33 mean square on both infit and outfit is low or moderate (L/M) fit (Wilson, 2005). We are regarding this as good fit from a clinical perspective, though we recognize that some would say that very low values, e.g., $< .75$ would be over-fitting. Infit or outfit values above 1.33 are regarded as high (H) or poor fitting patterns.

Therefore, in Figure 7, L/M on infit and L/M on outfit would be a pattern that is consistent with Rasch model expectations, i.e., good fit, and 4,451 persons, i.e., 83.4%, were in this fit group. The solid lines represent the actual item (dichotomous 0/1 categories) response means for each fit group, and the dashed lines represent the item means over all persons. The red vertical lines indicate the difference between fit group mean and the overall group mean.

In Figure 8, the L/M infit and high (H) outfit group consists of 572 people (10.7%) (termed Atypical Type 1). Relative to the overall mean, Atypical Type 1 response patterns on the are lower on the less severe symptoms, such as *FeelTired*, *TroubleSleeping*, *FeelSadTenseAngry*, and *ForgetProblemsRemembering* slightly lower than expected for mid-level severity symptoms, such as *ShakyHands*, *SweatHeartRaceGooseBumps*, *RunnyNoseEyesWater*, and *MoveTalkSlower* and only slightly higher than expected for high severity symptoms such as *UseDrugStopWithdrawal*, *SeeFeelHeardThingsNotReal*, *Diarrhea*, *Fever*, and *ConvulsionsSeizures*. The overall score may underestimate the severity of this group's withdrawal symptoms.

In Figure 9, the H infit and L/M outfit group consists of only 49 people (0.9%) (termed Atypical Type 2). Relative to the overall mean, Atypical Type 2 response patterns on the CWS are slightly lower than expected on the low severity symptoms described above but markedly higher than expected on most of the moderate and high severity symptoms. The overall score may underestimate the severity of this group's withdrawal symptoms.

In Figure 10, the H infit and H outfit group consists of 268 people (5%) (termed Atypical Type 3). Relative to the overall mean, Atypical Type 3 response patterns on the CWS are only slightly lower than expected or as expected for the low severity symptoms, but their response patterns are markedly higher than expected on the higher severity symptoms. The overall score may underestimate the severity of this group's withdrawal symptoms.

In summary, the CWS underestimates severity in all three atypical groups. Relative to what was expected based on their CWS measures, people in the Atypical Type 1 group tended to score lower than

expected mostly due to their tendency not to endorse low severity items. While people in the Atypical Types 2 and 3 groups tended not to endorse the low severity items, they tended to endorse the high severity items much more than expected. We recommend flagging the Atypical Groups 1, 2, and 3 for clinicians because these groups will tend to have scores that underestimate their current withdrawal.

Figure 6. CWS Pie Chart Highlighting the Fit Group Categories for the Subset of 5340 Persons Who Reported at Least One Withdrawal Symptom (Total N = 18, 113)

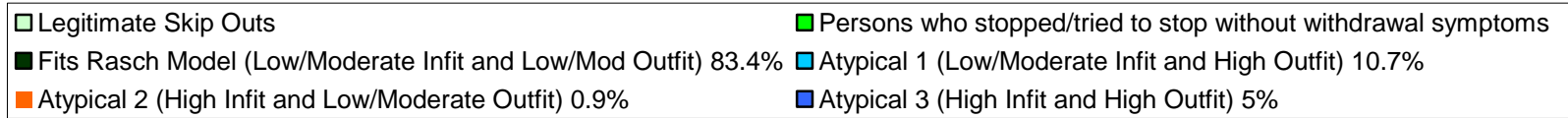
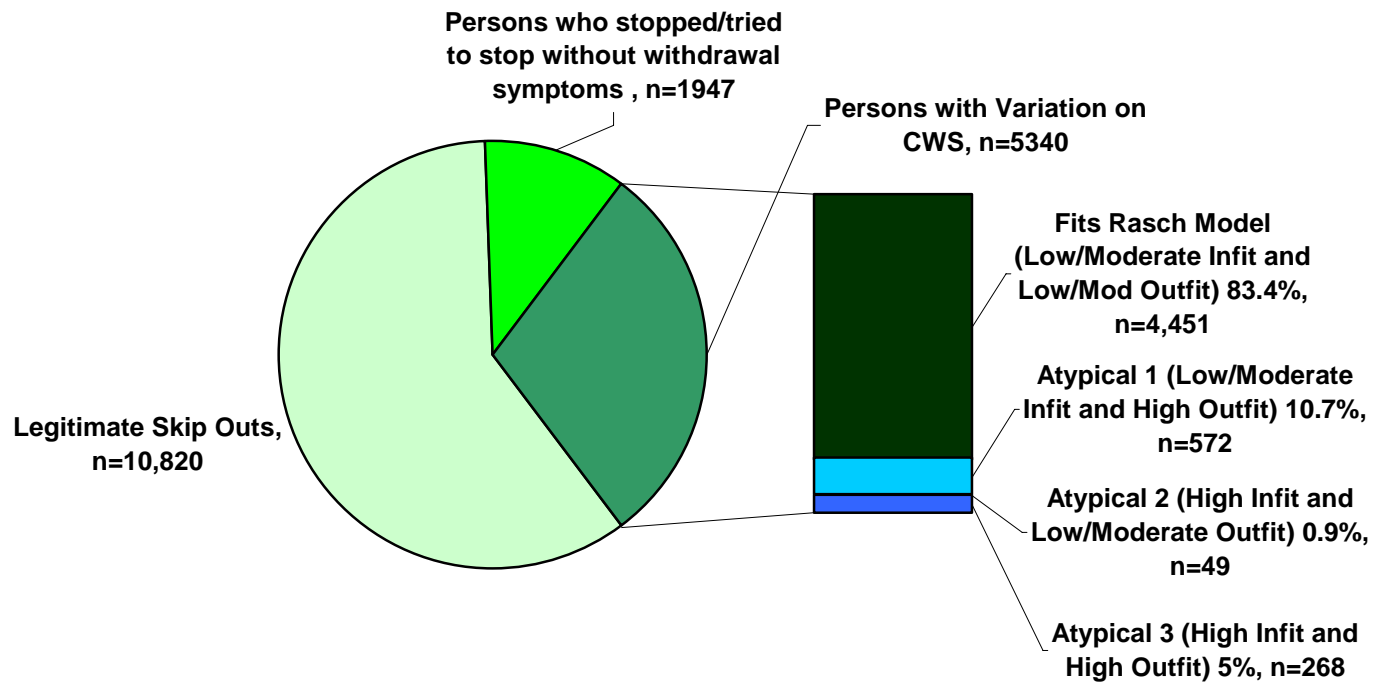


Figure 7

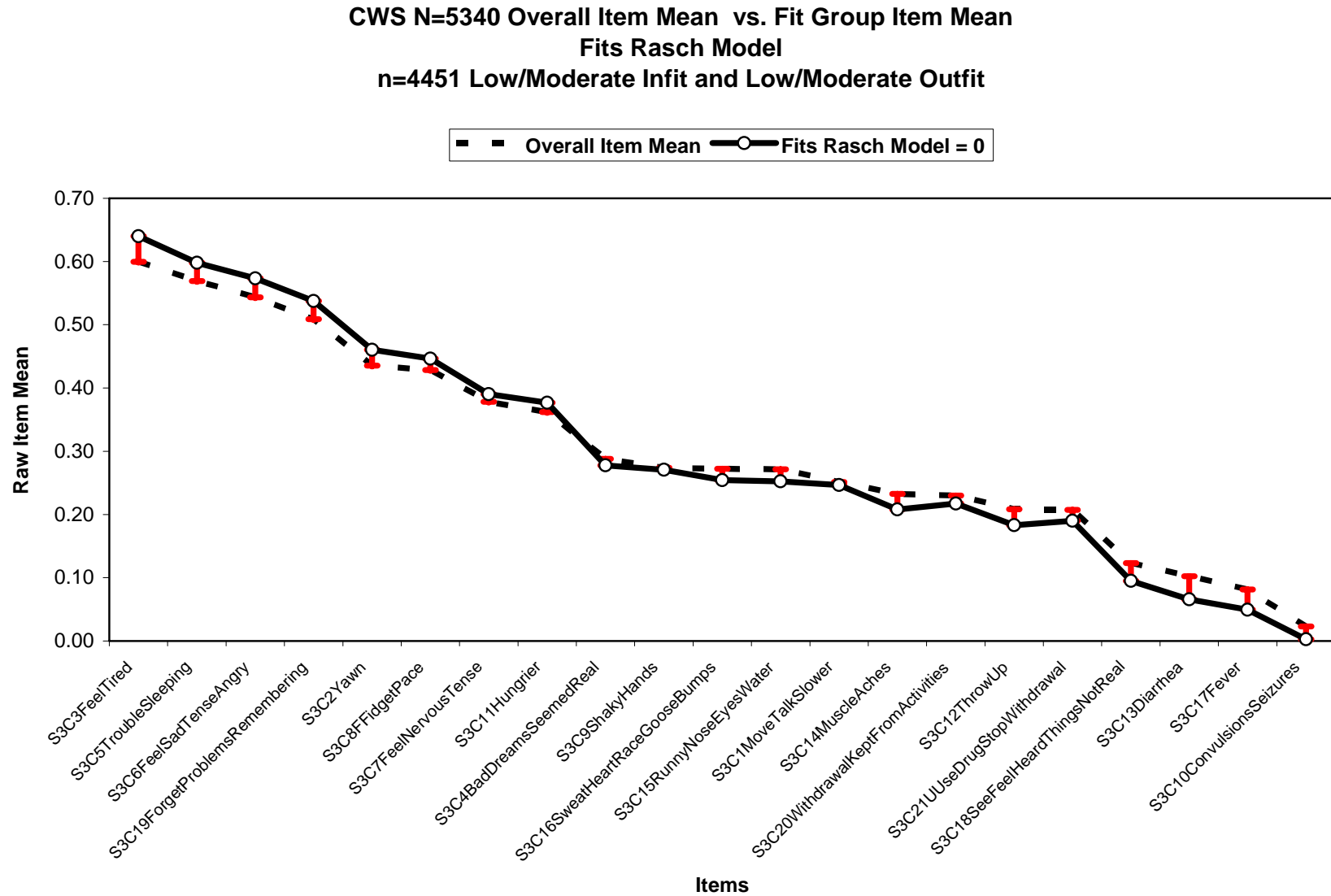


Figure 8

CWS N=5340 Overall Item Mean vs. Fit Group Item Mean
Atypical 1
n=572; 10.7%, Low/Moderate Infit and High Outfit

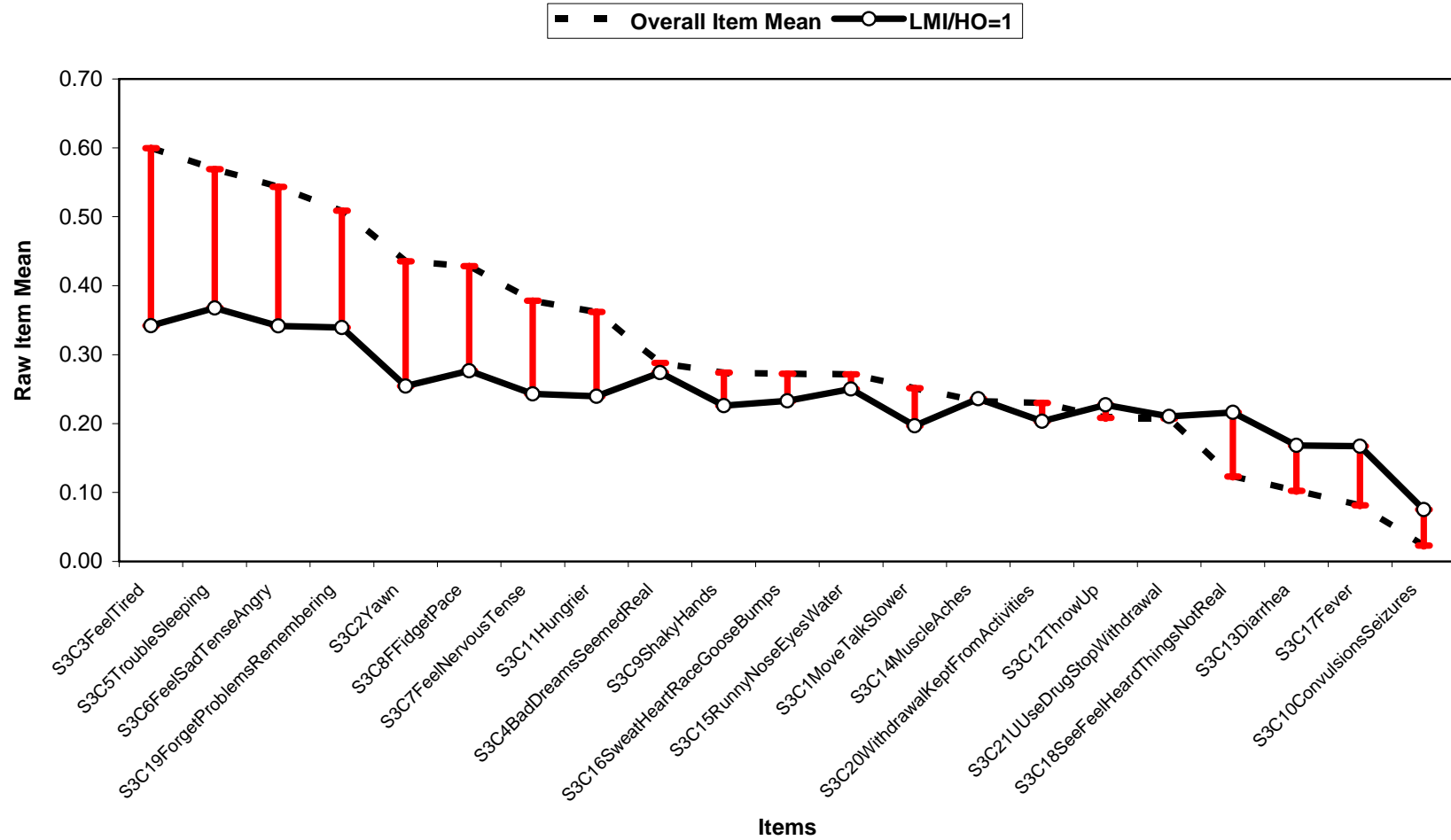


Figure 9

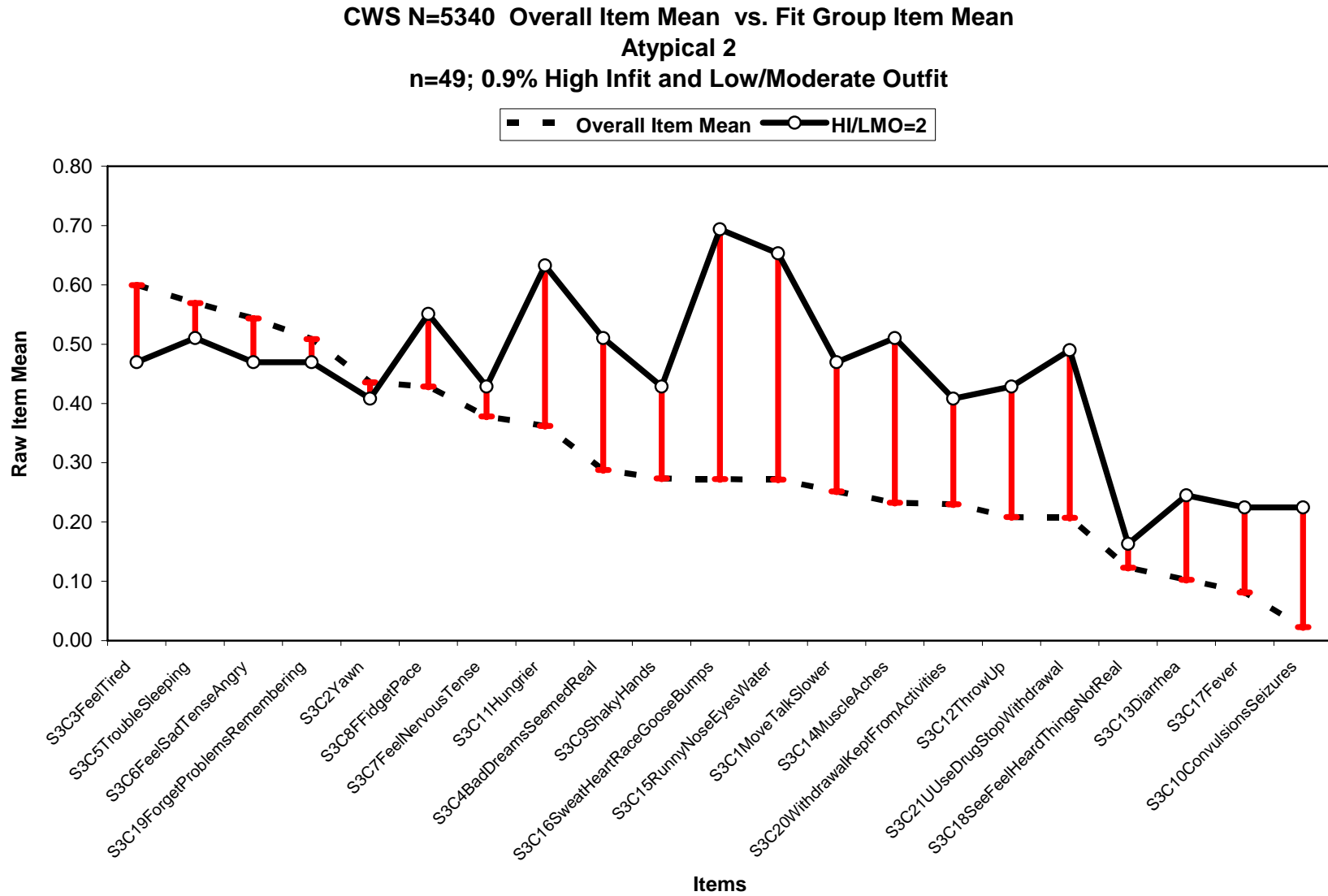
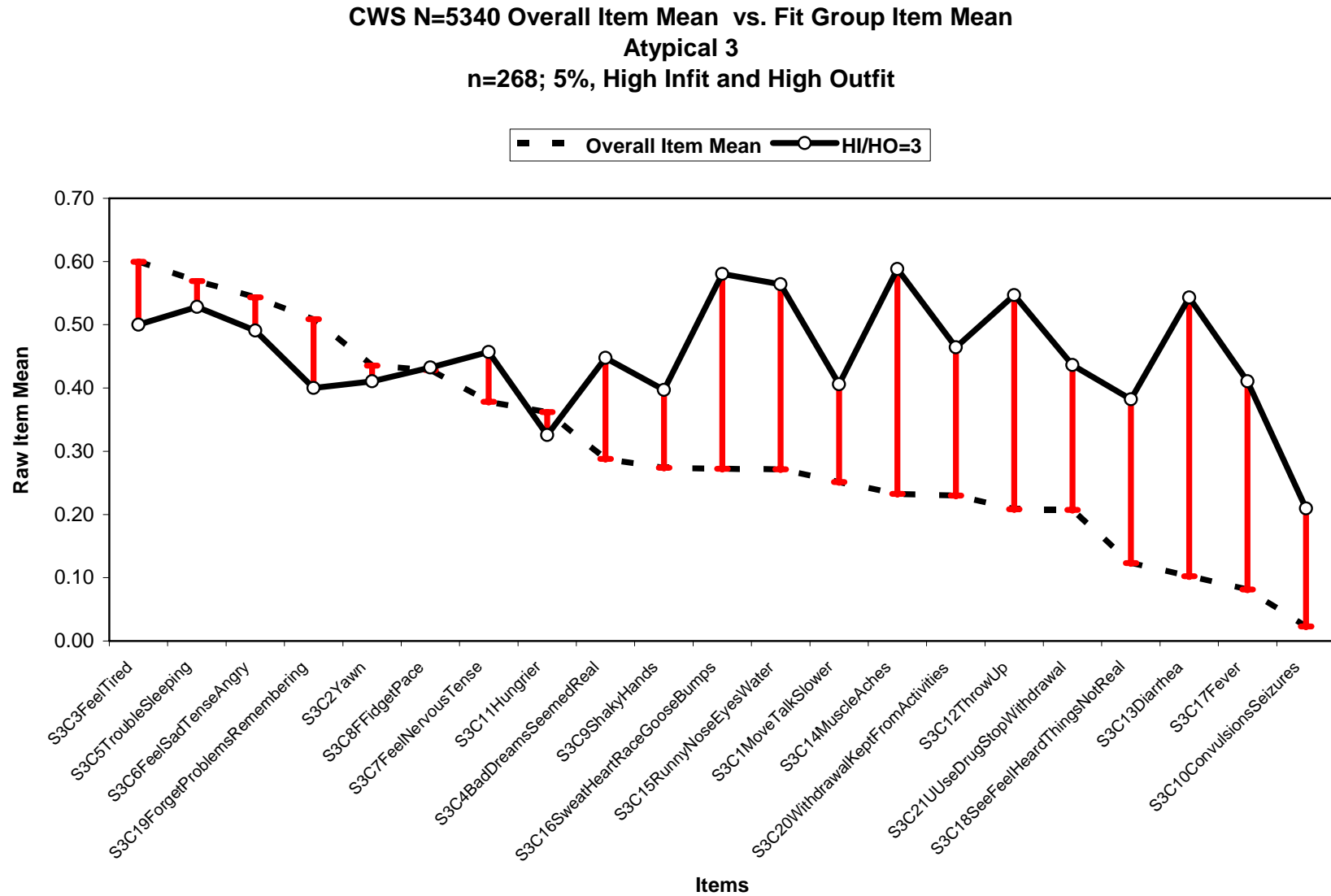


Figure 10



Recommended Actions and Deliberations

The CWS is useful in assessing its target construct. It has strong psychometric properties (see abstract and text). Additionally, we expected that the withdrawal symptoms should vary across drugs and therefore expected DIF across drug type. We observed this for opioids vs. all other drugs. This means that the construct of withdrawal may have a very different meaning for those who report opioid use as their primary drug than it does for persons who report other drugs as their primary drug.

There are several recommended actions and deliberations that might improve the CWS:

- To more completely cover the domain hierarchy, items could be added to the upper and lower ends of the continuum.
- One item, *S3C11Hungrier*, misfit using the criterion of .75-1.33 MNSQ for both infit and outfit.
- The results suggest an insignificant impact of DIF for age, and race because of the small number of items involved and the lack of theoretical implications for those items. There was no DIF for gender. However, further consideration should be given to the interpretation of the CWS for opioid users.
- Relative to what was expected based on their CWS measures, people in the Atypical Type 1 group tended to score lower than expected mostly due to their tendency not to endorse low severity items. While people in the Atypical Types 2 and 3 groups tended not to endorse the low severity items, they tended to endorse the high severity items much more than expected. We recommend flagging the Atypical Groups 1, 2, and 3 for clinicians because these groups will tend to have scores that underestimate their current withdrawal.
- Recommendation and Referral Summary (GRRS): We recommend flagging these three groups for clinicians in the relevant section of the *GAIN*.
 - Relative to the overall mean, Atypical Type 1 response patterns are lower than expected (i.e., lower than group average) on the less severe symptoms, such as *S3C3_ "Feel tired"*; *S3C5_ "Trouble sleeping"*; *S3C6_ "Feel sad, tense, or angry"*; and *S3C19_ "Forget a lot of things or have problems remembering"* and slightly lower than expected for mid-level severity symptoms, such as *S3C9_ "Have shaky hands"*; *S3C16_ "Sweat more than usual, have your heart race, or goose bumps"*; *S3C15_ "Runny nose or eyes watering more than usual"*; and *S3C1_ "Move and talk much slower than usual"*; and only slightly higher than expected for high severity symptoms, such as *S3C21_ "Use the same or other drug to stop or avoid withdrawal symptoms"*; *S3C18_ "See, feel, or hear things that are not real"*; *S3C13_ "Diarrhea"*; *S3C17_ "Fever"*; and *S3C10_ "Convulsions or seizures."* The overall score may underestimate the severity of this group's withdrawal symptoms.
 - Relative to the overall mean, Atypical Type 2 response patterns on the CWS are slightly lower than expected on the low severity symptoms described above but markedly higher than expected on most of the moderate and high severity symptoms.
 - Relative to the overall mean, Atypical Type 3 response patterns on the CWS are only slightly lower than expected or as expected for the low severity symptoms, but their response patterns are markedly higher than expected on the higher severity symptoms.
- More work on construct validity would be helpful to understand these fit groups better and to ensure proper interpretation of measures.

References

- American Psychiatric Association (2000). *Diagnostic and Statistical Manual (DSM-IV-TR, 4th Ed.* Washington, DC: American Psychiatric Association.
- Bond, T.G. & Fox, C.M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences. (2nd Ed.)* Mahwah, NJ: Erlbaum Associates.
- Conrad, K.J. & Smith, E.V. (2004). International conference on objective measurement: Applications of Rasch analysis in health care. *Medical Care, 2004; 42* (suppl I) 1-6.
- Dennis, M. L., Chan, Y-F., & Funk, R. R.(2006). Development and validation of the GAIN short screener (GSS) for internalizing, externalizing, and substance use disorders and crime/violence problems among adolescents and adults. *The American Journal on Addictions, 15*, 80-91.
- Dennis, M. L. Conrad, K. J. & Chan, Y. (2008). Variation in *DSM-IV* symptom severity depending on type of drug. Manuscript in progress.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, N.J.: Erlbaum.
- Linacre, J.M. (1998). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement, 2*(3), 266-283.
- Linacre J.M. (2006). Data Variance Explained by Measures, *Rasch Measurement Transactions, 20:1* p. 1045.
- Linacre, J.M. (2009). *Winsteps Rasch Measurement (Version 3.68).* www.winsteps.com. Author.
- Norman, G.R., Sloan, J.A., & Wyrwich, K.W. (2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Medical Care, 41*(5), 582-592.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Danmarks Paedagogiske Institut. (Republished Chicago: The University of Chicago Press: 1980).
- Reckase, M. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*, 207-230.
- Smith, E.V. (2001). Evidence for the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. *Journal of Applied Measurement, 2*(3), 281-311.
- Smith, E.V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement, 3*(2), 205-231.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach.* Mahwah, NJ: Erlbaum Associates.
- Wright, B.D., & Stone, M.H. (1979). *Best test design.* Chicago: University of Chicago, MESA Press.

ACKNOWLEDGEMENT: This development of this paper was supported by the Center for Substance Abuse Treatment (CSAT), Substance Abuse and Mental Health Services Administration (SAMHSA) via Westat under contract 270-07-0191 to Dr. Dennis at Chestnut Health Systems in Bloomington, Illinois using data provided by the following grants and contracts from CSAT (TI-11320, TI-11317, TI-11321, TI-11323, TI-11324, TI-11422, TI-11424, TI-11423, TI-11894, TI-11874, TI-11888, TI-11892, TI-11871, TI-13309, TI-13356, TI-13305, TI-13340, TI-13344, TI-13322, TI-13323, TI-13345, TI-13308, TI-13354, TI-13313, TI-14254, TI-14376, TI-14311, TI-14196, TI-14214, TI-14261, TI-14090, TI-14189, TI-14252, TI-14283, TI-14355, TI-14272, TI-14103, TI-14267, TI-14315, TI-14188, TI-14271, TI-15686, TI-15671, TI-15486, TI-15545, TI-15672, TI-15475, TI-15678, TI-15447, TI-15461, TI-15433, TI-15481, TI-15514, TI-15478, TI-15413, TI-15483, TI-15670, TI-15674, TI-15479, TI-15682, TI-15467, TI-15511, TI-15562, TI-13601, TI-13190, TI-12541, TI-00567, TI-15415, TI-15421, TI-15438, TI-15446, TI-15458, TI-15466, TI-15469, TI-15485, TI-15489, TI-15524, TI-15527, TI-15577, TI-15584, TI-15586, TI-15677, TI-16386, TI-16400, TI-16414, TI-16904, TI-16915, TI-16928, TI-16939, TI-16961, TI-16984, TI-16992, TI-17046, TI-17055, TI-17070, TI-17071, TI-17334, TI-17433, TI-17434, TI-17446, TI-17475, TI-17476, TI-17484, TI-17490, TI-17517, TI-17523, TI-17530, TI-17604, TI-17605, TI-17638, TI-17728, TI-17755, TI-17761, TI-17763, TI-17765, TI-17769, TI-17779, TI-17786, TI-17788, TI-17812, TI-17825, TI-17830, TI-18406; Contract 207-98-7047, Contract 277-00-6500, Contract 270-2003-00006), the National Institute on Alcohol Abuse and Alcoholism (NIAAA) (R01 AA 10368), the National Institute on Drug Abuse (NIDA) (R37 DA11323; R01 DA 018183), the Illinois Criminal Justice Information Authority (95-DB-VX-0017), and the Illinois Office of Alcoholism and Substance Abuse (PI 00567). The opinions are those of the author and do not reflect official positions of the contributing project directors or government.

011210

Appendix. Crosswalk of DSM-IV-TR Drug Withdrawal Symptoms and Primary Drug Crossed with GAIN Current Withdrawal Scale Items Listed in Rasch Difficulty Order

DSM-IV-TR symptom ^a	CWS Items	Rasch Item Hierarchy: Difficulty Order - Easiest to Hardest to Endorse	Alcohol	Amphetamine	Cocaine	Marijuana	Opioids
Fatigue	3. Feel tired	1		B1	B1		
Insomnia or hypersomnia	5. Have trouble sleeping, including sleeping too much or not being able to sleep	2	B3	B3	B3		B9
Dysphoric mood	6. Feel sad, tense, or angry	3					B1
General symptom of withdrawal	19. Forget a lot of things or have problems remembering	4	GEN	GEN	GEN	GEN	GEN
Yawning	2. Yawn more than usual	5					B7
Psychomotor agitation	8. Fidget, pace, wring your hands, or have trouble sitting still	6	B6	B5	B5		
Anxiety	7. Feel really nervous	7	B7				
Increased appetite	11. Feel hungrier than usual	8		B4	B4		
Vivid, unpleasant dreams	4. Have bad dreams that seemed real	9		B2	B2		
Hand tremors	9. Have shaky hands	10	B2				
Autonomic hyperactivity; pupillary dilation, piloerections, or sweating	16. Sweat more than usual, have your heart race or get goose bumps	11	B1				B5
Lacrimation or rhinorrhea	15. Have a runny nose or eyes watering more than usual	12					B4
Psychomotor retardation	1. Move and talk much slower than usual	13		B5	B5		

DSM-IV-TR symptom ^a	CWS Items	Rasch Item Hierarchy: Difficulty Order - Easiest to Hardest to Endorse	Alcohol	Amphetamine	Cocaine	Marijuana	Opioids
Muscle aches	14. Have muscle aches	14					B3
Criteria c	20. Have any of these withdrawal problems kept you from doing social, family, job, or other activities	15	C	C	C	C	C
Nausea or vomiting	12. Throw up or feel like throwing up	16	B4				B2
General symptom of withdrawal	21. Have you used the same or another drug to stop or avoid having any of these withdrawal symptoms?	17	GEN	GEN	GEN	GEN	GEN
Transient visual, tactile, or auditory hallucinations or illusions	18. See, feel, or hear things that are not real	18	B5				
Diarrhea	13. Have diarrhea	19					B6
Fever	17. Have a fever	20					B8
Grand mal seizures	10. Have convulsions or seizures	21	B8				

^a American Psychiatric Association (2000). *Diagnostic and Statistical Manual (DSM-IV-TR, 4th Ed.)*. DSM. Criteria: B = the substance-specific syndrome causes clinically significant distress or impairment in social, occupational, or other important areas of functioning. C = the symptoms are not due to a general medical condition and are not better accounted for by another mental disorder. GEN = General symptom.