

Validation of the HIV Scale to the Rasch Measurement Model,  
GAIN Methods Report 1.1

Kendon J. Conrad  
University of Illinois at Chicago

Karen M. Conrad  
University of Illinois at Chicago

Michael L. Dennis  
Chestnut Health Systems

Barth B. Riley  
Chestnut Health Systems

Rod Funk  
Chestnut Health Systems

**Abstract**

**Purpose.** The purpose of this report is to provide a brief psychometric analysis of the *HIV Risk Scale*, using the Rasch measurement model. The 35-item *HIV Risk Scale* consists of 31 yes/no items and four recency items related to the risk of contracting HIV.

**Methods.** Data were analyzed on 6269 persons who presented for substance problem screening. Rasch analysis included an examination of: person and item reliabilities; dimensionality; 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.

**Results.** The *HIV Risk Scale* performs well as a measure of the construct of HIV risk. Both items and scales form the theoretically expected hierarchies with a Rasch person internal consistency reliability of .79 and an item reliability of 1.00. Cronbach's alpha is .86. The persons' responses generally conformed to the expectations of the Rasch model. Of the 35 items in the *HIV Risk Scale*, a DIF > .5 SD occurred in 10 items for males vs. females, 3 items for youth vs. adults, 12 items for race when using African American as the reference group, and 15 items for primary substances when using opiates as the referent and in 5 items involving needle use when alcohol was the referent. For the above contrasts, 22 of the 35 items exhibited DIF for one or more groups. While such DIF usually cancels out in the overall scale, it should be taken into account when interpreting items, for tailoring interventions to these groups, and in some cases corrected when the goal is to compare these groups. In terms of person fit, 81% of the sample exhibited person infit and outfit that was low or moderate (L/M) and are thus, regarded as fitting the Rasch model expectations well from a clinical perspective. Five percent (5%) of the persons had a L/M infit and high (HI) outfit and are referred to as Atypical Type 1, where the overall score may underestimate severity since these tended to be persons engaging in high risk sex who are not reporting lower risk victimization. Six percent (6%) of the persons had HI infit and L/M outfit (called Atypical Type 2) and tended to be valid low scorers who endorsed a few victimization items. The remaining 8% of persons had HI infit and HI outfit (called Atypical Type 3) and tended to be valid high scorers who also had an overall score that may underestimate severity slightly and tended to be needle users with less risky sex than expected.

**Citation.** Conrad, K. J., Conrad, K. M., Dennis, M. L., Riley, B.B., & Funk (2009). Validation of the HIV Scale to the Rasch Measurement Model, GAIN Methods Report 1.1. Chicago, IL: Chestnut Health Systems. Retrieved from

[http://www.chestnut.org/li/gain/psychometric\\_reports/Conrad\\_et\\_al\\_2009\\_HIV\\_Rasch\\_Report.pdf](http://www.chestnut.org/li/gain/psychometric_reports/Conrad_et_al_2009_HIV_Rasch_Report.pdf)

## Purpose

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

## Background

The *HIV Risk Scale* is a scale within the larger *Global Appraisal of Individual Needs (GAIN;* Dennis et al 2003) which is a series of related instruments that share the same general instructions, questions, scoring, interpretation, and clinical decision trees. 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 *HIV Risk Scale* is a count of past-year symptoms related to HIV risk.

## HIV Risk Subscales and Items

The *HIV Risk Scale* consists of three subscales with a total of 35 items. The response format is Yes/No (coded: no=0, yes=1). Its subscales are the *General Victimization Scale (GVS;* 15 items; time frame: ever except for four recency items), the *Needle Problem Scale (NPS;* 9 items; time frame: past year), and the *Sexual Risk Scale (S<sub>x</sub>RS;* 11 items; time period: past year). The *GVS* is a count of types of victimization experienced by the respondent (including physical, emotional and sexual victimization) and the number of traumagenic factors involved in the victimization (including origin and duration, type and relation of perpetrator, etc). The *NPS* is a count of the number of endorsed activities related to using a needle to inject drugs that increase the risk of transmission of bloodborne diseases). While the *NPS* items are the most difficult (rare) items, they are only asked if there is past year needle use and represent little respondent burden. The *S<sub>x</sub>RS* is a count of the number of endorsed sexual activities that might put the client at risk for sexually transmitted diseases or other difficulties. The original *S<sub>x</sub>RS* consisted of 12 items. In the revised 11-item scale, the items, *SexWMale* and *SexWFemale* were recoded into a new variable, *ANYSEX*, that measures any sexual activity in the past year. The scale names, item stems, *GAIN* item numbers, Rasch output item location codes, and item labels are shown in Table 1 below.

**Table 1. Scale and Item Information**

Subscale Name and Item Stem	GAIN Item Number	Item Label	Rasch Output Item Number
<b>General Victimization Scale</b> Time Period is past 12 months except for recency items, #12-#15.			
Has anyone <b>ever</b> : ...(Response set Yes=1, No=0)			
1. Attacked you with a gun, knife, stick, bottle or other weapon?	E9A	AttackUW/Weapn	1
2. Hurt you by striking or beating you to the point that you had bruises, cuts, or broken bones or otherwise physically abused you?	E9B	HurtUStrike/Beat	2
3. Pressured or forced you to participate in sexual acts against your will, including your regular sexual partner, a family member or friend?	E9C	ForceSexAgainstUWill	3
4. Abused you emotionally; that is, did or said things to make you feel very bad about yourself or your life?	E9D	AbusedUEmotionly	4
5. About how old were you the first time any of these things happened to you? (E9e_d - Scored as 1 IF E9e UNDER 18) E9e18. [Is E9e under 18?]	E9E	Age1stAbuse	5
Did any of the above things happen ... (Response set Yes=1, No=0)			
6. Several times or over a long period of time?	E9F	SevralTimesOrLongTime	6
7. With more than one person involved in hurting you?	E9G	MorThan1PersonHurtU	7
8. Where one or more of the people involved was a family member close family friend, professional or someone else you had trusted?	E9H	FamilyOrTrusted1	8
9. Where you were afraid for your life or that you might be seriously injured?	E9J	Araid4ULife/Injury	9
10. And result in oral, vaginal or anal sex?	E9K	OralVaginalAnalSex	10
11. And people you told did not believe or help you?	E9M	PplNotBelieve/HelpU	11
Are you currently worried that someone might ... (Response set Yes=1, No=0)			
12. Attack you with a gun, knife, stick, bottle or other weapon?	E9N	CrntWoryAttackUWeapn	12
13. Hurt you by striking or beating or otherwise physically abuse you?	E9P	CrntWoryStrike/BeatU	13
14. Pressure or force you to participate in sexual acts against your will?	E9Q	CrntWoryForceSex	14
15. Abuse you emotionally?	E9R	CrntWoryAbuseEmotn	15
<b>Needle Problem Scale</b>			
During the past 12 months, did you... ...(Response set Yes=1, No=0)			
1. Use a needle to shoot up drugs?	R1A	UseNdleShootUpDrugs	16

Subscale Name and Item Stem	GAIN Item Number	Item Label	Rasch Output Item Number
2. Reuse a needle that you had used before?	R1B	ReuseNdleUUsedBefor	17
3. Reuse a needle without cleaning it with bleach or boiling water	R1C	ReuseNdleW/OClean	18
4. Use a needle that you knew or suspected someone else had used before?	R1D	UseNdleSomeoneUsed	19
5. Use someone else's rinse water, cooker, or cotton after they did?	R1E	UseSome1WaterEtc	20
6. Ever skip cleaning your needle with bleach or boiling water after you were done?	R1F	SkipCleanNdleAfterU	21
7. Let someone else use a needle after you used it?	R1G	Let1UseNdleAfterU	22
8. Let someone else use the rinse water, cooker or cotton after you did?	R1H	Let1UseH2OEtcAfterU	23
9. Allow someone else to inject you with drugs?	R1J	Allow1InjectUW/Drugs	24
<b>Sexual Risk Scale</b>			
During the past 12 months, did you... ... (Response set Yes=1, No=0)			
1. Have sex while you or your partner was high on alcohol or drugs?	R2A	SexWhenPartnerHighAOD	25.
2. Have sex with someone who was an injection drug user?	R2B	SexW/InjectnDrugUser	26.
3. Have sex involving anal intercourse (penis to butt)?	R2C	SexAnalIntercourse	27.
4. Have sex with a man who might have had sex with other men?	R2D	SexW/ManSexW/OthMan	28.
5. Trade sex to get drugs, gifts or money?	R2E	TradeSex4DrugGftMoney	29.
6. Use drugs, gifts, or money to purchase or get sex?	R2F	UseDrugs\$EtcToBuySex	30.
7. Have two or more sexual partners?	R2G	Two/MoreSexPartner12Mo	31.
<del>8. Have sex with a man? (Recoded to 13 below)</del>	<del>R2H</del>	<del>SexW/Male</del>	<del>32.</del>
<del>9. Have sex with a woman? (Recoded to 13 below)</del>	<del>R2J</del>	<del>SexW/Female</del>	<del>33.</del>
10. Have sex without using any kind of condom, dental dam or other barrier to protect you and your partner from diseases or pregnancy?	R2K	SexW/OBarrier	34.
11. Have a lot of pain during sex or after having had sex?	R2M	PainW/Sex	35.
12. Use alcohol or drugs to make sex last longer or hurt less?	R2N	UseAOD4LongSexHrtLes	36.
13. Any sexual activity? (RECODE OF R2H AND R2J into no ( had not engaged in sexual activity) or yes (engaged in sexual activity).	NA	ANYSEX	37.

### Data Source

Data on the 6269 cases reported in this paper came from 12 projects/programs including 64 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 *HIV Risk Scale* and the full GAIN (Dennis et al 2003) 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 are considered acceptable (Wilson, 2005). Low fit values, i.e., <.75, provide less motivation for data editing than do high values (Wilson, 2005, p. 129), unless obvious duplication is found, e.g., a repeated question or a double-scanned response form. Low fit values do not disturb the meaning of a measure. They merely reduce precision (Linacre & Wright, 1994). We used fit statistics to examine items that should be dropped. Person fit statistics were examined in order to inform better the clinical interpretation of the HIV measures. In summary:

- Persons having both infit and outfit that is low or moderate (L/M) are regarded as fitting well from a clinical perspective, so their measures are interpretable in the normal way, i.e., low measures indicate less and high indicate more of the construct .
- Persons having L/M infit and HI outfit are referred to as Atypical Type 1.
- Persons having HI infit and L/M outfit are called Atypical Type 2.
- Persons having HI infit and HI outfit are called Atypical Type 3.

The interpretation of these four groups, which we call “fit groups,” is unique to the HIV. In fact, although these four groups can be generated for any measure, the interpretation must always be tailored to the particular measure.

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 an overview and 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 of 6269 persons was predominantly under 18 years of age (69.5%) and male (65.7%). Almost half were Caucasian (44.4%), a quarter were African American (26.7%), and the remainder Hispanic or mixed race. Of the top five most severe primary drugs reported, marijuana was reported by 45% of the sample. The drug least often reported was opiates at 5.7%. Other drugs reported included amphetamines (12.9%), cocaine (12.2%), and alcohol (22%). Two percent of the sample reported other primary drugs.

**Table 2. Demographic Characteristics of the Sample (N=6269<sup>a</sup>)**

	Percent	Number
<b>Age, Mean (sd) 20.6 (9.5)</b>		
< 18 years	69.5	4357
≥18 years	30.4	1903
<b>Gender</b>		
Male	65.7	4118
Female	34.1	2140
<b>Race</b>		
African American	26.7	1673
Caucasian	44.4	2782
Hispanic	10.1	635
Mixed/other	18.2	1138
<b>Drug, primary, most severe</b>		
Alcohol	22.0	1380
Amphetamines	12.9	806
Marijuana	45.0	2819
Cocaine	12.2	764
Opiates	5.7	357
Other drug	2.0	124
<sup>a</sup> Numbers may not add up to 100% due to missing values		

**Table 3. Person and Item Reliability**

SUMMARY OF 5624 MEASURED (NON-EXTREME) PERSONS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	7.2	34.6	-2.60	.63	.97	.0	.86	.0
S.D.	4.6	1.8	1.51	.19	.40	1.1	1.23	.9
MAX.	30.0	35.0	2.93	1.39	3.41	5.5	9.90	6.8
MIN.	1.0	11.0	-5.57	.45	.20	-3.2	.04	-1.8
REAL RMSE	.70	ADJ.SD	1.34	SEPARATION	1.92	PERSON RELIABILITY	.79	
MODEL RMSE	.66	ADJ.SD	1.36	SEPARATION	2.05	PERSON RELIABILITY	.81	
S.E. OF PERSON MEAN = .02								

- Person reliability is acceptable at .79.
- Persons are placed reliably on the ruler.
- Cronbach alpha =.86.
- Cronbach's alpha is higher because it estimates extreme scores as measured perfectly, i.e., with no error.
- A separation value of 1.92 gives approximately two separation levels, thus splitting the persons into about 3 groups on the Rasch ruler (Figure 1).

MINIMUM EXTREME SCORE: 534 PERSONS  
DELETED: 111 PERSONS  
VALID RESPONSES: 98.9%

SUMMARY OF 6158 MEASURED (EXTREME AND NON-EXTREME) PERSONS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	6.6	34.6	-2.98	.75				
S.D.	4.8	1.9	1.90	.41				
MAX.	30.0	35.0	2.93	2.05				
MIN.	.0	4.0	-7.07	.45				
REAL RMSE	.88	ADJ.SD	1.69	SEPARATION	1.92	PERSON RELIABILITY	.79	
MODEL RMSE	.85	ADJ.SD	1.70	SEPARATION	1.99	PERSON RELIABILITY	.80	
S.E. OF PERSON MEAN = .02								

- Good item reliability of 1.00.
- Item separation is high at 30.81 meaning items are placed reliably on the ruler.

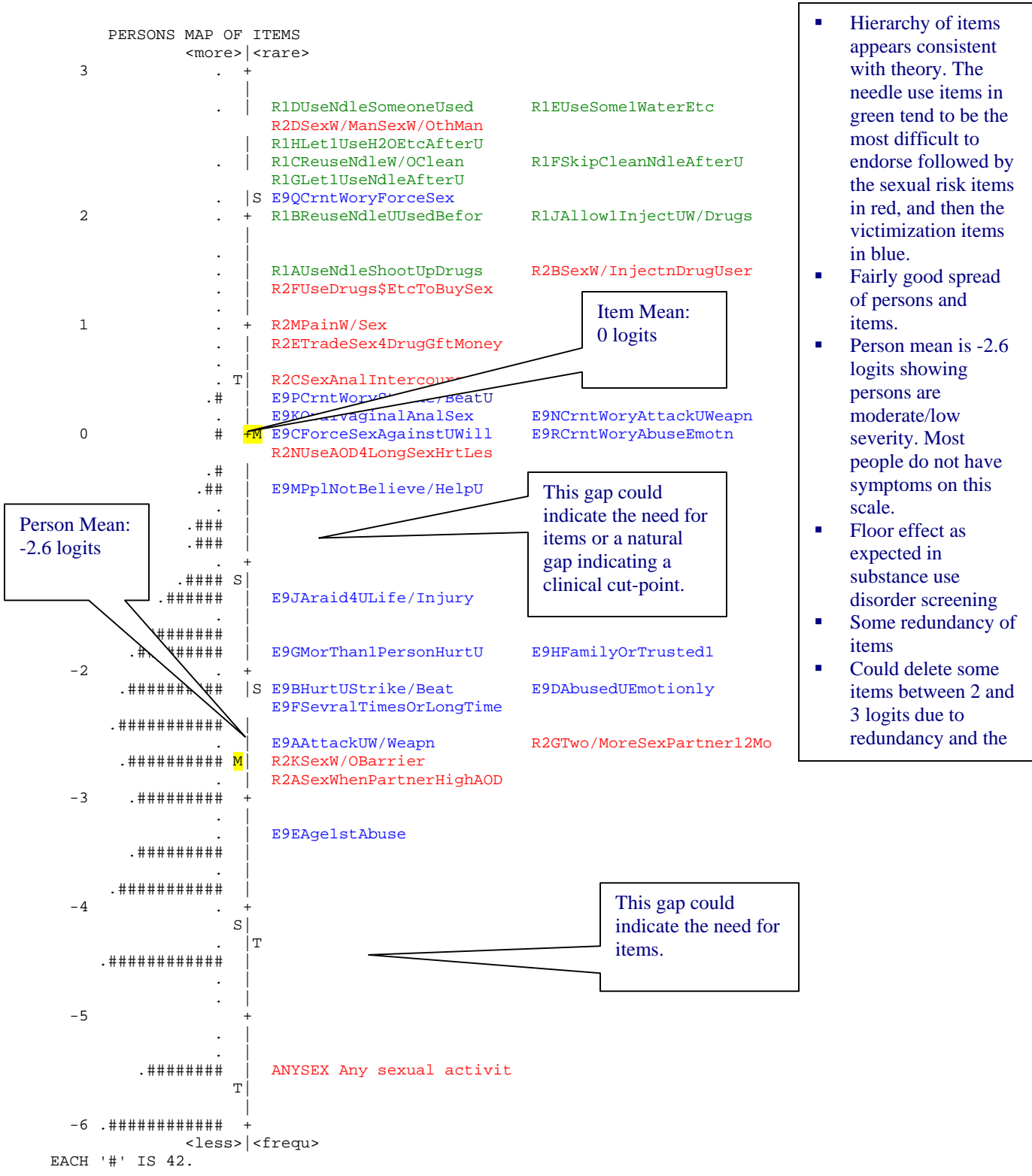
PERSON RAW SCORE-TO-MEASURE CORRELATION = .94 (approximate due to missing data)  
CRONBACH ALPHA (KR-20) PERSON RAW SCORE RELIABILITY = .86 (approximate due to missing data)

SUMMARY OF 35 MEASURED (NON-EXTREME) ITEMS

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	1152.8	5564.4	.00	.06	.98	-.3	.89	-1.3
S.D.	1243.3	37.0	2.13	.03	.14	4.7	.40	5.1
MAX.	4982.0	5616.0	2.68	.12	1.42	9.9	2.02	9.9
MIN.	79.0	5502.0	-5.58	.03	.82	-9.9	.22	-9.0
REAL RMSE	.07	ADJ.SD	2.13	SEPARATION	30.81	ITEM RELIABILITY	1.00	
MODEL RMSE	.07	ADJ.SD	2.13	SEPARATION	31.15	ITEM RELIABILITY	1.00	
S.E. OF ITEM MEAN = .37								

DELETED: 2 ITEMS  
UMEAN=.000 USCALE=1.000  
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.97 (approximate due to missing data)  
194755 DATA POINTS. APPROXIMATE LOG-LIKELIHOOD CHI-SQUARE: 107525.91

**ITEM STATISTICS**  
**Figure 1. Wright Item Map**



- Hierarchy of items appears consistent with theory. The needle use items in green tend to be the most difficult to endorse followed by the sexual risk items in red, and then the victimization items in blue.
- Fairly good spread of persons and items.
- Person mean is -2.6 logits showing persons are moderate/low severity. Most people do not have symptoms on this scale.
- Floor effect as expected in substance use disorder screening
- Some redundancy of items
- Could delete some items between 2 and 3 logits due to redundancy and the

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

CONTRAST 1 FROM PRINCIPAL COMPONENT ANALYSIS OF  
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --	Modeled
Total raw variance in observations	=	66.4	100.0%
Raw variance explained by measures	=	31.4	47.3%
Raw variance explained by persons	=	7.1	10.6%
Raw Variance explained by items	=	24.3	36.6%
Raw unexplained variance (total)	=	35.0	52.7%
Unexplned variance in 1st contrast	=	4.5	6.7%

STANDARDIZED RESIDUAL LOADINGS FOR ITEMS (SORTED BY LOADING)

CON-TRAST	LOADING	MEASURE	INFIT	OUTFIT	ENTRY	NUMBER	ITEM
1	.73	2.06	.83	.44	A	17	R1BReuseNdleUUsedBefor
1	.68	2.27	.82	.31	B	21	R1FSkipCleanNdleAfterU
1	.66	2.41	.82	.55	C	18	R1CReuseNdleW/OClean
1	.65	2.62	.84	.31	D	20	R1EUseSomelWaterEtc
1	.65	2.44	.84	.26	E	23	R1HLet1UseH2OetcAfterU
1	.64	2.40	.84	.31	F	22	R1GLet1UseNdleAfterU
1	.62	2.68	.82	.22	G	19	R1DUseNdleSomeoneUsed
1	.62	2.04	.87	.90	H	24	R1JAllow1InjectUW/Drugs
1	.61	1.53	.86	.95	I	16	R1AUseNdleShootUpDrugs
1	.25	1.57	.91	.75	J	26	R2BSexW/InjectnDrugUser
1	.12	-2.84	1.13	1.13	K	25	R2ASexWhenPartnerHighAOD
1	.11	-2.63	1.23	1.27	L	34	R2KSexW/OBarrier
1	.08	-5.58	1.42	2.02	M	37	ANYSEX Any sexual activity - r
1	.02	-2.48	1.28	1.42	N	31	R2GTwo/MoreSexPartner12Mo
1	.02	-.01	1.07	1.08	O	36	R2NUseAOD4LongSexHrtLes
1	.00	2.65	.99	.81	P	28	R2DSexW/ManSexW/OthMan
1	-.28	-2.08	.85	.76	a	6	E9FSevralTimesOrLongTime
1	-.25	-3.36	.84	1.01	b	5	E9EAgelstAbuse
1	-.23	-1.32	.92	.78	c	9	E9JAraid4ULife/Injury
1	-.22	-2.21	.91	.83	d	4	E9DAbusedUEmotionly
1	-.22	-.41	1.01	.76	e	11	E9MPplNotBelieve/HelpU
1	-.21	-1.86	.96	.85	f	8	E9HFamilyOrTrusted1
1	-.19	-2.18	.93	.89	g	2	E9BHurtUStrike/Beat
1	-.19	-1.88	1.00	.93	h	7	E9GMorThan1PersonHurtU
1	-.18	-2.49	1.03	1.13	i	1	E9AAttackUW/Weapon
1	-.18	-.03	1.03	.94	j	15	E9RCrntWoryAbuseEmotn
1	-.16	-.05	.90	.69	k	3	E9CForceSexAgainstUWill
1	-.15	.16	.88	.57	l	10	E9KOralVaginalAnalSex
1	-.13	.34	1.06	1.05	m	13	E9PCrntWoryStrike/BeatU
1	-.10	.19	1.13	1.58	n	12	E9NCrntWoryAttackUWeapon
1	-.08	2.18	1.01	.71	o	14	E9QCrntWoryForceSex
1	-.03	.91	.95	.92	p	29	R2ETradeSex4DrugGftMoney
1	-.02	1.05	1.09	1.39	q	35	R2MPainW/Sex
1	-.01	.52	1.14	1.37	r	27	R2CSEXAnalIntercourse
1	.00	1.42	1.09	1.30	Q	30	R2FUseDrugs\$EtcToBuySex

- 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,  $\geq 30\%$  is considered a moderate measurement dimension, and  $\geq 20\%$  is considered a minimal measurement dimension. The 20% criterion is taken from Reckase (1979).
- 47.3% of the variance is explained by the items, i.e., a strong principal measurement dimension
- 12.7% of the unexplained variance is explained by the 1<sup>st</sup> residual contrast. Thus, the variance explained by the first factor of residuals supports unidimensionality.
- Examination of the loadings indicates that the 1<sup>st</sup> residual contrast represents the dimensions of “needle use” which tends to have the highest loadings vs. “sexual risk” and “general victimization” combined which tend to have moderate and low loadings.
- As we noted, this is not a strong dimension, i.e., 1<sup>st</sup> contrast. Additionally, high vs. low severity items will usually tend to form separate dimensions even though they are, conceptually, on the same dimension.
- The items are listed with higher loadings appearing first.
- The loadings closest to zero are the sexual risk items.
- The loadings with high negatives are the victimization items.
- Therefore, we see the three conceptual areas corresponding to three areas of loadings. In spite of this they do not form separate dimensions but contribute to the general dimension of HIV Risk at distinct levels of severity.
- The PCA was conducted using *Winsteps v. 3.68*

**Table 5. Item Misfit Order in Terms of INFIT and OUTFIT Mean Square (MNSQ)**

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S. E.	INFIT		OUTFIT		PTMEA CORR.	EXACT OBS%	MATCH EXP%	ITEM
					MNSQ	ZSTD	MNSQ	ZSTD				
37	4982	5544	-5.58	.05	1.42	9.9	2.02	9.2	A .50	89.9	89.9	ANYSEX Any sexual activity
12	627	5549	.19	.05	1.13	4.1	1.58	5.6	B .29	88.3	89.6	E9NCrntWoryAttackUWeapn
31	2654	5537	-2.48	.03	1.28	9.9	1.42	9.9	C .45	64.8	73.5	R2GTwo/MoreSexPartner12Mo
35	318	5536	1.05	.06	1.09	1.8	1.39	3.2	D .23	94.0	94.5	R2MPainW/Sex
27	479	5538	.52	.05	1.14	3.6	1.37	3.5	E .26	90.9	91.9	R2CSEXAnalIntercourse
30	235	5538	1.42	.07	1.09	1.5	1.30	2.3	F .21	95.6	95.9	R2FUseDrugs\$EtcToBuySex
34	2800	5538	-2.63	.03	1.23	9.9	1.27	9.3	G .48	66.6	73.7	R2KSexW/OBarrier
25	2983	5502	-2.84	.03	1.13	8.8	1.13	4.4	H .54	68.8	74.2	R2ASexWhenPartnerHighAOD
1	2708	5608	-2.49	.03	1.03	2.3	1.13	4.6	I .55	72.5	73.5	E9AAttackUW/Weapn
36	706	5536	-.01	.05	1.07	2.5	1.08	1.0	J .33	87.6	88.4	R2NUseAOD4LongSexHrtLes
13	563	5548	.34	.05	1.06	1.8	1.05	.6	K .32	89.8	90.6	E9PCrntWoryStrike/BeatU
15	728	5551	-.03	.04	1.03	1.1	.94	-.8	L .37	87.0	88.1	E9RCrntWoryAbuseEmotn
5	3481	5511	-3.36	.03	.84	-9.9	1.01	.5	M .66	83.1	76.3	E9EAge1stAbuse
14	126	5548	2.18	.09	1.01	.1	.71	-2.3	N .20	97.7	97.8	E9QCcrntWoryForceSex
11	933	5546	-.41	.04	1.01	.2	.76	-4.0	O .42	84.2	85.2	E9MPplNotBelieve/HelpU
7	2098	5560	-1.88	.03	1.00	-.2	.93	-2.3	P .54	73.4	74.5	E9GMorThan1PersonHurtU
28	79	5521	2.65	.12	.99	-.1	.81	-1.4	Q .16	98.6	98.6	R2DSEXW/ManSexW/OthMan
8	2086	5565	-1.86	.03	.96	-3.1	.85	-5.0	R .55	75.1	74.5	E9HFamilyOrTrusted1
16	221	5612	1.53	.07	.86	-2.5	.95	-.4	q .28	96.5	96.2	R1AUseNdleShootUpDrugs
29	354	5537	.91	.06	.95	-1.1	.92	-.7	p .30	93.8	93.9	R2ETradeSex4DrugGftMoney
2	2405	5610	-2.18	.03	.93	-4.8	.89	-3.9	o .58	75.1	73.6	E9BHurtUStrike/Beat
9	1601	5564	-1.32	.03	.92	-5.0	.78	-5.9	n .52	79.2	77.8	E9JAraid4ULife/Injury
4	2434	5601	-2.21	.03	.91	-6.5	.83	-6.3	m .59	76.2	73.6	E9DAbusedUEmotionly
26	204	5504	1.57	.08	.91	-1.5	.75	-2.1	l .26	96.6	96.4	R2BSEXW/InjectnDrugUser
24	142	5533	2.04	.09	.87	-1.8	.90	-.7	k .24	97.5	97.5	R1JAllow1InjectUW/Drugs
3	744	5605	-.05	.04	.90	-3.9	.69	-4.5	j .42	88.5	88.0	E9CForceSexAgainstUWill
10	635	5553	.16	.05	.88	-4.0	.57	-5.7	i .41	89.8	89.5	E9KOralVaginalAnalSex
6	2290	5553	-2.08	.03	.85	-9.9	.76	-9.0	h .61	78.2	73.9	E9FSevraltimesOrLongTime
22	104	5613	2.40	.10	.84	-1.8	.31	-6.9	g .24	98.2	98.2	R1GLet1UseNdleAfterU
23	100	5616	2.44	.11	.84	-1.9	.26	-7.6	f .24	98.3	98.2	R1HLet1UseH2OetcAfterU
20	85	5616	2.62	.11	.84	-1.7	.31	-6.7	e .22	98.5	98.5	R1EUseSome1WaterEtc
17	141	5614	2.06	.09	.83	-2.3	.44	-5.3	d .26	97.7	97.5	R1BReuseNdleUUsedBefor
21	117	5616	2.27	.10	.82	-2.2	.31	-7.0	c .25	98.1	97.9	R1FSkipCleanNdleAfterU
19	81	5616	2.68	.12	.82	-1.8	.22	-8.2	b .23	98.6	98.6	R1DUseNdleSomeoneUsed
18	103	5616	2.41	.10	.82	-2.1	.55	-3.8	a .24	98.3	98.2	R1CReuseNdleW/OClean
MEAN	1152.8	5564.4	.00	.06	.98	-.3	.89	-1.3		87.6	87.9	
S.D.	1243.3	37.0	2.13	.03	.14	4.7	.40	5.1		10.4	9.8	

- Here we see that being sexually active (ANYSEX) is the most misfitting item in terms of both INFIT and OUTFIT.
- It is also the easiest item.
- As we will see later, this is caused by persons who use needles, a high severity symptom, who do not have sex. Since ANYSEX is such an easy item, this causes misfit, i.e., easy items are expected to be endorsed by persons with higher scores, such as needle users.

- 4 of the 5 items with high OUTFIT are sexual risk items. This again reflects higher scorers who, unexpectedly, are not having sex.
- Items with low outfit tend to be needle use items. This means they are overly predictive of HIV risk (beyond chance). However, the infit is within acceptable range which indicates that the item is performing appropriately.

**Table 6. Most Unexpected Responses in Terms of Measure**

MOST UNEXPECTED RESPONSES

ITEM	MEASURE	PERSON
		4311132 5255432 4 325443 5544333221 544332221
		55306749872719484757999959483649629037755540054457
		92175246229631421071378384695910304856523782145045
		03665727401606782983847718394929669652837427664766
		-----high-----
37 ANYSEX Any se	-5.58 A	0000.....00.....0.....
5 E9EAgelstAbus	-3.36 M	.0.....
25 R2ASexWhenPar	-2.84 H	0000.....
34 R2KSexW/OBarr	-2.63 G	0000.....
31 R2GTwo/MoreSe	-2.48 C	0000.....1.....
2 E9BHurtUStrik	-2.18 o	.0.....
7 E9GMorThan1Pe	-1.88 P	.00.....
3 E9CForceSexAg	-.05 j	.....1.....
15 E9RCrntWoryAb	-.03 L	.....1.....
10 E9KOralVagina	.16 i	.....1.....
12 E9NCrntWoryAt	.19 B	.....1.111.11..
13 E9PCrntWorySt	.34 K	.....1.....
27 R2CSEXAnalInt	.52 E	...111.11.1.1.1.....
29 R2ETradeSex4D	.91 p	.....1.....
35 R2MPainW/Sex	1.05 D	.....1.....1.....11.11..11.....
30 R2FUseDrugs\$E	1.42 F	.....1.....1.1..1..1..1..1.....
16 R1AUseNdleSho	1.53 q	.11...1.. .1...1....1... ..11.....1
26 R2BSEXW/Injec	1.57 l	.....1.....
24 R1JAllow1Inje	2.04 k	.1...1...1.....11.....1.1.....
17 R1BReuseNdleU	2.06 d	111.....1.....1.....
14 E9QCrntWoryFo	2.18 N	...1 .....1.1.....
21 R1FSkipCleanN	2.27 c	111.....
22 R1GLet1UseNdl	2.40 g	111...1.....
18 R1CReuseNdleW	2.41 a	111.....11.....
23 R1HLet1UseH2O	2.44 f	111.....1.....
20 R1EUseSome1Wa	2.62 e	111.....1.....
28 R2DSEXW/ManSe	2.65 Q	...111.11.1.111.....
19	2.68 b	1.1.....
		-----low-----
		43111329852554324453254439455443332217755443322217
		55306746272719481777999954683649629036525540054455
		92175247429631422081378388395910304852833782145046
		0366572 0160678 9 384771 9492966965 742766476

- This table is another view of item misfit. It shows the most misfitting items in terms of item measures.
- The items and their measures are shown on the left and the persons with the most unexpected responses are on the right.

- For example, Person 4590 had a relatively high measure of 0.28 (person mean = -2.60), but unexpectedly did not endorse a number of easier sex items, such as *AnySex*. See also in Table 10.
- This is an excellent example of why ANYSEX misfits.

**Table 7. Most Misfitting Response Strings in Terms of Outfit Mean Square (OUTMNSQ)**

MOST MISFITTING RESPONSE STRINGS		PERSON										
ITEM	OUTMNSQ											
		4311132	5255432	4	325443	5544333221	544332221					
		55306749872719484	757999959483649629037755540054457									
		92175246229631421	071378384695910304856523782145045									
		03665727401606782	983847718394929669652837427664766									
	high	-----										
37 ANYSEX Any se	2.02	A	0000	.....	00	.....	0	.....				
12 E9NCrntWoryAt	1.58	B	....	.....					1	111	11	
31 R2GTwo/MoreSe	1.42	C	0000	.....	..	1	.....					
35 R2MPainW/Sex	1.39	D	.....	1	.....	1	.....	11	11	11		
27 R2CSEXAnalInt	1.37	E	....	111	11	1	1	1	.....			
30 R2FUseDrugs\$E	1.30	F	.....	1	.....	1	1	1	1	1	1	
34 R2KSexW/OBarr	1.27	G	0000	.....								
25 R2ASexWhenPar	1.13	H	0000	.....								
13 E9PCrntWorySt	1.05	K	....	.....						1	.....	
15 E9RCrntWoryAb	.94	L	....	.....							1	
5 E9EAgelstAbus	1.01	M	.0	.....								
14 E9QCrntWoryFo	.71	N	...1	.....	1	1	.....					
7 E9GMorThan1Pe	.93	P	.00	.....								
28 R2DSEXW/ManSe	.81	Q	....	111	11	1	111	.....				
16 R1AUseNdleSho	.95	q	.11	....	1	..	1	1	1	1	1	
29 R2ETradeSex4D	.92	p	.....	.....	1	.....			1	.....		
2 E9BHurtUStrik	.89	o	..0	.....								
26 R2BSEXW/Injec	.75	l	.....	.....	1	.....						
24 R1JAllow1Inje	.90	k	..1	....	1	1	.....	11	.....	1	1	
3 E9CForceSexAg	.69	j	.....	.....						1	.....	
			-----low-----									
		43111329852554324453254439455443332217755443322217										
		55306746272719481777999954683649629036525540054455										
		92175247429631422081378388395910304852833782145046										
		0366572 0160678 9 384771 9492966965 742766476										

- 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.
- ANYSEX has the highest outfit (2.02) because a few people with high overall measures did not endorse engaging in any sexual activity in the past year (the zeroes) even though it was a low severity item (-5.58). These were persons at high risk such as Person 4590 had a high score of .28 but did not endorse sex and Person 412 who endorsed needle use but did not endorse sex. Table 9 shows these persons and their measures.

**Table 8. Item Entry Order Showing Infit, Outfit, and Point-by-Measure (a.k.a. item/total) Correlations**

ITEM STATISTICS: ENTRY ORDER

ENTRY NUMBER	RAW		MEASURE	MODEL S.E.	INFIT		OUTFIT		PTMEA CORR.	EXACT OBS%	MATCH EXP%	ITEM
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD				
1	2708	5608	-2.49	.03	1.03	2.3	1.13	4.6	.55	72.5	73.5	E9AAttackUW/Weapn
2	2405	5610	-2.18	.03	.93	-4.8	.89	-3.9	.58	75.1	73.6	E9BHurtUStrike/Beat
3	744	5605	-.05	.04	.90	-3.9	.69	-4.5	.42	88.5	88.0	E9CForceSexAgainstUWill
4	2434	5601	-2.21	.03	.91	-6.5	.83	-6.3	.59	76.2	73.6	E9DAbusedUEmotionly
5	3481	5511	-3.36	.03	.84	-9.9	1.01	.5	.66	83.1	76.3	E9EAgelstAbuse
6	2290	5553	-2.08	.03	.85	-9.9	.76	-9.0	.61	78.2	73.9	E9FSevralTimesOrLongTime
7	2098	5560	-1.88	.03	1.00	-.2	.93	-2.3	.54	73.4	74.5	E9GMorThan1PersonHurtU
8	2086	5565	-1.86	.03	.96	-3.1	.85	-5.0	.55	75.1	74.5	E9HFamilyOrTrusted1
9	1601	5564	-1.32	.03	.92	-5.0	.78	-5.9	.52	79.2	77.8	E9JAraid4ULife/Injury
10	635	5553	.16	.05	.88	-4.0	.57	-5.7	.41	89.8	89.5	E9KOralVaginalAnalSex
11	933	5546	-.41	.04	1.01	.2	.76	-4.0	.42	84.2	85.2	E9MPplNotBelieve/HelpU
12	627	5549	.19	.05	1.13	4.1	1.58	5.6	.29	88.3	89.6	E9NCrntWoryAttackUWeapn
13	563	5548	.34	.05	1.06	1.8	1.05	.6	.32	89.8	90.6	E9PCrntWoryStrike/BeatU
14	126	5548	2.18	.09	1.01	.1	.71	-2.3	.20	97.7	97.8	E9QCrntWoryForceSex
15	728	5551	-.03	.04	1.03	1.1	.94	-.8	.37	87.0	88.1	E9RCrntWoryAbuseEmotn
16	221	5612	1.53	.07	.86	-2.5	.95	-.4	.28	96.5	96.2	R1AUseNdleShootUpDrugs
17	141	5614	2.06	.09	.83	-2.3	.44	-5.3	.26	97.7	97.5	R1BReuseNdleUUsedBefor
18	103	5616	2.41	.10	.82	-2.1	.55	-3.8	.24	98.3	98.2	R1CReuseNdleW/OClean
19	81	5616	2.68	.12	.82	-1.8	.22	-8.2	.23	98.6	98.6	R1DUseNdleSomeoneUsed
20	85	5616	2.62	.11	.84	-1.7	.31	-6.7	.22	98.5	98.5	R1EUseSomelWaterEtc
21	117	5616	2.27	.10	.82	-2.2	.31	-7.0	.25	98.1	97.9	R1FSkipCleanNdleAfterU
22	104	5613	2.40	.10	.84	-1.8	.31	-6.9	.24	98.2	98.2	R1GLet1UseNdleAfterU
23	100	5616	2.44	.11	.84	-1.9	.26	-7.6	.24	98.3	98.2	R1HLet1UseH2OetcAfterU
24	142	5533	2.04	.09	.87	-1.8	.90	-.7	.24	97.5	97.5	R1JAllow1InjectUW/Drugs
25	2983	5502	-2.84	.03	1.13	8.8	1.13	4.4	.54	68.8	74.2	R2ASexWhenPartnerHighAOD
26	204	5504	1.57	.08	.91	-1.5	.75	-2.1	.26	96.6	96.4	R2BSexW/InjectnDrugUser
27	479	5538	.52	.05	1.14	3.6	1.37	3.5	.26	90.9	91.9	R2CSEXAnalIntercourse
28	79	5521	2.65	.12	.99	-.1	.81	-1.4	.16	98.6	98.6	R2DSEXW/ManSexW/OthMan
29	354	5537	.91	.06	.95	-1.1	.92	-.7	.30	93.8	93.9	R2ETradeSex4DrugGftMoney
30	235	5538	1.42	.07	1.09	1.5	1.30	2.3	.21	95.6	95.9	R2FUseDrugs\$EtcToBuySex
31	2654	5537	-2.48	.03	1.28	9.9	1.42	9.9	.45	64.8	73.5	R2GTwo/MoreSexPartner12Mo
32	DELETED											R2HSEXW/Male
33	DELETED											R2JSEXW/Female
34	2800	5538	-2.63	.03	1.23	9.9	1.27	9.3	.48	66.6	73.7	R2KSEXW/OBarrier
35	318	5536	1.05	.06	1.09	1.8	1.39	3.2	.23	94.0	94.5	R2MPainW/Sex
36	706	5536	-.01	.05	1.07	2.5	1.08	1.0	.33	87.6	88.4	R2NUseAOD4LongSexHrtLes
37	4982	5544	-5.58	.05	1.42	9.9	2.02	9.2	.50	89.9	89.9	ANYSEX Any sexual activity - recode of r2h and r2j
MEAN	1152.8	5564.4	.00	.06	.98	-.3	.89	-1.3		87.6	87.9	
S.D.	1243.3	37.0	2.13	.03	.14	4.7	.40	5.1		10.4	9.8	

- This table includes point-by-measure correlations (a.k.a., item/total). The correlations among the needle use items and the total measure tend to be low, in the .20s. This seems to be due to the combination of the high severity of these items along with the fact that they form a somewhat distinct aspect of HIV risk.

**PERSON STATISTICS**

**Table 9. Persons with Most Unexpected Responses in Terms of Measures**

MOST UNEXPECTED RESPONSES		PERSON		MEASURE		ITEM	
						3 23 3 1 131112233122112212221	
						75541142678913560237950664741283089	
						high	
4590	112223511421	.28	U	0.00	0.0	.....	1.11111.1
3523	221115011525	-.15	R	0000	0.0...0	.....	1..1.11111..
1076	222221501621	-.38		0.00	0.0	.....	.....1.....
1316	114211701525	-.38	J	0.00	0.0.0.0	.....	1.11.11111.1
1655	222221701111	-2.67		.....	.....	.....	1.....1.
824	112221811321	-2.72		.....	.....	.....	.....1.....
967	112221701221	-2.72	I	.....	.....	.....	.....1.1.....
2442	114211810321	-2.72		.....	.....	.....	.....1.1.....
3727	111114810421	-2.72		.....	.....	.....	.....1.1.....
5720	111111701111	-2.84		.....	.....	.....	.....1.....
2291	222221400321	-3.05		.....	.....	.....	.....1.....
412	112221701525	-3.06	A	0.	.....	.....	.....1.....
2828	112222011321	-3.06		.....	.....	.....	.....1.....
3447	221113310421	-3.06	Q	.....	.....	.....	.....1.....
4916	112224211421	-3.06	Y	.....	.....	.....	.....1.....
5130	224211301221	-3.06		.....	.....	.....	.....1.1.....
5766	112221700321	-3.06	Z	.....	.....	.....	.....1.....
4709	221114811421	-3.35		0.	.....	.....	.....1.....
578	222221701321	-3.44		.....	.....	.....	.....1.....
3713	221113810421	-3.81		.....	.....	.....	.....11.....
2938	221111400321	-3.91	N	.....	.....	.....	.....1.....
948	112221701321	-3.92	H	0.	.....	.....	.....11.....
3581	111114111421	-3.92		.....	.....	.....	.....1.....
4937	222222110221	-3.92		.....	.....	.....	.....1.....
4987	112221701621	-3.92		.....	.....	.....	.....1.....
5974	111111300321	-3.92		.....	.....	.....	.....1.....
463	114211600525	-4.55	B	.....	.....	.....	.....1.1.....
523	112221811221	-4.55	C	.....	.....	.....	.....1.....
758	113211601321	-4.55	E	.....	.....	.....	.....1.....
762	113211501421	-4.55	F	.....	.....	.....	.....1.....
1355	112221701421	-4.55	L	.....	.....	.....	.....1.1.....
2086	112221501321	-4.55		.....	.....	.....	.....1.....
2949	223211501321	-4.55		.....	.....	.....	.....1.....
3206	111113610421	-4.55	P	.....	.....	.....	.....1.....
3636	221112910111	-4.55		.....	.....	.....	.....1.....
3909	221112711421	-4.55		.....	.....	.....	.....1.....
4412	222221601111	-4.55		.....	.....	.....	.....1.....
4699	112225011111	-4.55	W	.....	.....	.....	.....1.....
5354	222221501321	-4.55		.....	.....	.....	.....1.....
5899	223211500221	-4.55		.....	.....	.....	.....1.....
756	113211701525	-5.57	D	.....	.....	.....	.....1.....

- This table illustrates misfitting persons in terms of item measures.
- Again, we see some of the relatively high scorers such as person #4590 with unexpected non-endorsement of low severity items, i.e., zeroes on the right.
- The other unexpected responses are the low scorers like Person #412 who endorsed a couple of high severity items like #16, *RIAUseNdleShootU Drugs* and #23, *RIHLetIUseH2OEtcAfterU*.
- Would these misfitting persons fit an unusual profile? We suggest they should be flagged and given special consideration even though they may have very low scores. These needle users with very low scores may still be at high risk.

**Table 10. Persons with Most Misfitting Response Strings in Terms of OUTMNSQ**

MOST MISFITTING RESPONSE STRINGS			ITEM
PERSON	OUTMNSQ		3 23 3 1 131112233122112212221
			75541142678913560237950664741283089
		high-----	
412	112221701525	9.90 A	0.....1.....1...
463	114211600525	9.90 B	..... .1 .1....
523	112221811221	9.90 C	.....1.....
756	113211701525	9.90 D	.....1.....
758	113211601321	9.90 E	.....1.....
762	113211501421	9.90 F	.....1.....
947	222221810221	9.90 G	.0.....1111.11.1.1
948	112221701321	9.90 H	0.....11...1....
967	112221701221	9.90 I	.....1.1...1....
1316	114211701525	9.90 J	0.00.0.0.0.....1.11.11111.1
1350	112221910525	9.90 K	00.....1.11.11111.1
1355	112221701421	9.90 L	.....1.1.....
2546	112221301321	9.90 M	.....1.....
2938	221111400321	9.90 N	.....1.....
2997	222221700321	9.90 O	.....1.1.1.1.1..
3206	111113610421	9.90 P	.....1.....
3447	221113310421	9.90 Q	.....1.....1..
3523	221115011525	9.90 R	0000.0...0.....1.1.11111..
4186	224212311525	9.90 S	.....1.11...11..
4388	112221600221	9.90 T	.....1.1.11....
4590	112223511421	9.90 U	0.00.0.....1.11111.1
4665	112223211421	9.90 V	.....1.1.1.1.1.1
4699	112225011111	9.90 W	.....1.....
6110	112221501525	9.90 X	.0..... . . .1111 .1...1
4916	112224211421	9.86 Y	.....1.....1..
5766	112221700321	9.86 Z	.....1.....1..
		low-----	
			35231342678913131112233122112212221
			7 54 1 1 560237950664741283089

- This table shows the persons who unexpectedly endorsed an item given their overall measure. Much of the person misfit is being caused by *Needle Problem Scale* items #16-23, when persons with low severity are unexpectedly endorsing the high severity items.
- For example, Person #412 one of the most misfitting persons with a low measure (-3.16 shown in previous table) but an unexpected endorsement of a number of the *Needle Problem* items.
- Person # 1316 is another example of a respondent who did not endorse some of the easier items but endorsed the needle items.

**Table 11. Summary of Category Structure**

CATEGORY LABEL	SCORE	OBSERVED COUNT	OBSVD %	SAMPLE AVRG	EXPECT	INFINIT MNSQ	OUTFIT MNSQ	COHERENCE M->C	ESTIM C->M	DISCR
0	0	154408	78	-3.44	-3.44	1.05	1.28	90%	94%	0
1	1	40347	20	.55	.55	.99	.79	73%	63%	1.00
MISSING		2085	1	-1.73						

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

M->C = Does Measure imply Category?

C->M = Does Category imply Measure?

78% of all responses were zeros (item not endorsed).

## Differential Item Functioning for Age, Gender, Race, and Primary Drug Severity for the HIV Risk 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 for the HIV Risk scale items is based on  $\geq 1.07$  logit difference for all comparisons which is approximately half a standard deviation (Table 3, bottom panel) for the items (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 *HIV Risk* items. Table 12 contains the data that formed these graphs, and provides the information to compute the contrasts between groups on each item. For example, to compute the DIF contrast between males and females on *E9CForceSexAgainstUWill* subtract the calibration for Female = -1.02 from the calibration for Male = 1.29 to yield a DIF contrast of 2.31.

*Gender DIF.* There were 12 items with significant DIF by gender using our criterion of half a standard deviation or  $\geq 0.96$  logit absolute difference. When we examine the chart, we see seven items (five with some of the largest contrasts) that were much easier for females to endorse: *E9DAbusedUEmotionly*, *E9HFamilyOrTrusted1*, *E9CForceSexAgainstUWill*, *E9KOralVaginalAnalSex*, *R2ETradeSex4DrugGftMoney*, *R2MPainW/Sex*, and *E9QCmntWoryForceSex*.

Five items were easier for males to endorse: *E9AAttackUW/Weapn*, *E9GMorThan1PersonHurtU*, *E9NCmntWoryAttackUWeapn*, *R2CSEXAnalIntercourse*, and *R2FUseDrugs\$EtcToBuySex*. All of these items with DIF concerned either victimization (variable name began with E) or sexual risk (variable name began with R2). In general, males were more likely to experience physical violence and females were more likely to experience sexual abuse.

While these items tend to balance out or cancel each other in the overall measure so that measures of males and females can be compared with each other, it is clear that the item calibrations differ substantially. In other words, males and females have very different item hierarchies. Therefore, using these calibrations would not be appropriate in a context such as computerized adaptive testing where only a subset of the items would be used. In such cases, the gender specific hierarchy should be used, i.e., separate measures and item calibrations for males and females.

*Age DIF.* There were 3 items with significant DIF by age. One was easier for youth to endorse: *9EAge1stAbuse*, but this is likely due to the simple fact that the youth group has an age ceiling of 18 which limits the possible age of 1<sup>st</sup> abuse. Of course, it would not be possible for youths, i.e., defined as those under 18, to endorse being  $\geq 18$  at first abuse. The two that were easier for adults to endorse were *R2ETradeSex4DrugGftMoney*, *R2FUseDrugs\$EtcToBuySex*. The adults more readily endorsed the items involving exchanges of money or commodities for sex. While these are indeed significant DIF contrasts, we do not see this as being problematic since the adult/youth co-calibration is appropriately high.

*Race DIF.* Since the largest DIF tended to occur between African Americans vs. Caucasians and Hispanics, we chose African Americans as the reference group and compared them to all other races. Only two items were easier for African Americans to endorse: *R2ETradeSex4DrugGftMoney*, *R2FUseDrugs\$EtcToBuySex*. Ten items were more difficult for African Americans to endorse: *E9EAge1stAbuse*, *R1AUseNdleShootUpDrugs*, *R1JAllow1InjectUW/Drugs*, *R1BReuseNdleUUsedBefor*,

*R1FSkipCleanNdleAfterU, R1GLet1UseNdleAfterU, R1CReuseNdleW/OClean, R1HLet1UseH2OetcAfterU, R1EUseSome1WaterEtc, and R1DUseNdleSomeoneUsed.* Based on this evidence, needle use is much rarer among African Americans than among the rest. For African Americans there is greater risk in the area of exchanging sex for drugs, money, and gifts.

*Primary Drug Severity DIF.* Although only 6% of the sample reported opiates as their primary drug, the most significant DIF contrasts tended to involve opiates, so we elected to examine these contrasts. Sixteen items exhibited DIF. Ten items were easier for the opiates group than the other primary drug groups to endorse: *R1AUseNdleShootUpDrugs, R2BSexW/InjectnDrugUser, R1JAllow1InjectUW/Drugs, R1BReuseNdleUUsedBefor, R1FSkipCleanNdleAfterU, R1GLet1UseNdleAfterU, R1CReuseNdleW/OClean, R1HLet1UseH2OetcAfterU, R1EUseSome1WaterEtc, and R1DUseNdleSomeoneUsed.* All of the items involved needle use. Six items were harder for the opiates group to endorse: *E9EAge1stAbuse, R2GTwo/MoreSexPartner12Mo, E9MPplNotBelieve/HelpU, E9RCrntWoryAbuseEmotn, E9NCrntWoryAttackUWeapn, and E9PCrntWoryStrike/BeatU.* Five of the 6 items involved general victimization and the remaining item involved sex partners. (Although not reported here, when using alcohol as the reference drug, the 7 items that involved DIF were all needle problem items and were harder for the alcohol group to endorse compared to the other primary drug groups).

Figure 2. HIV Gender DIF

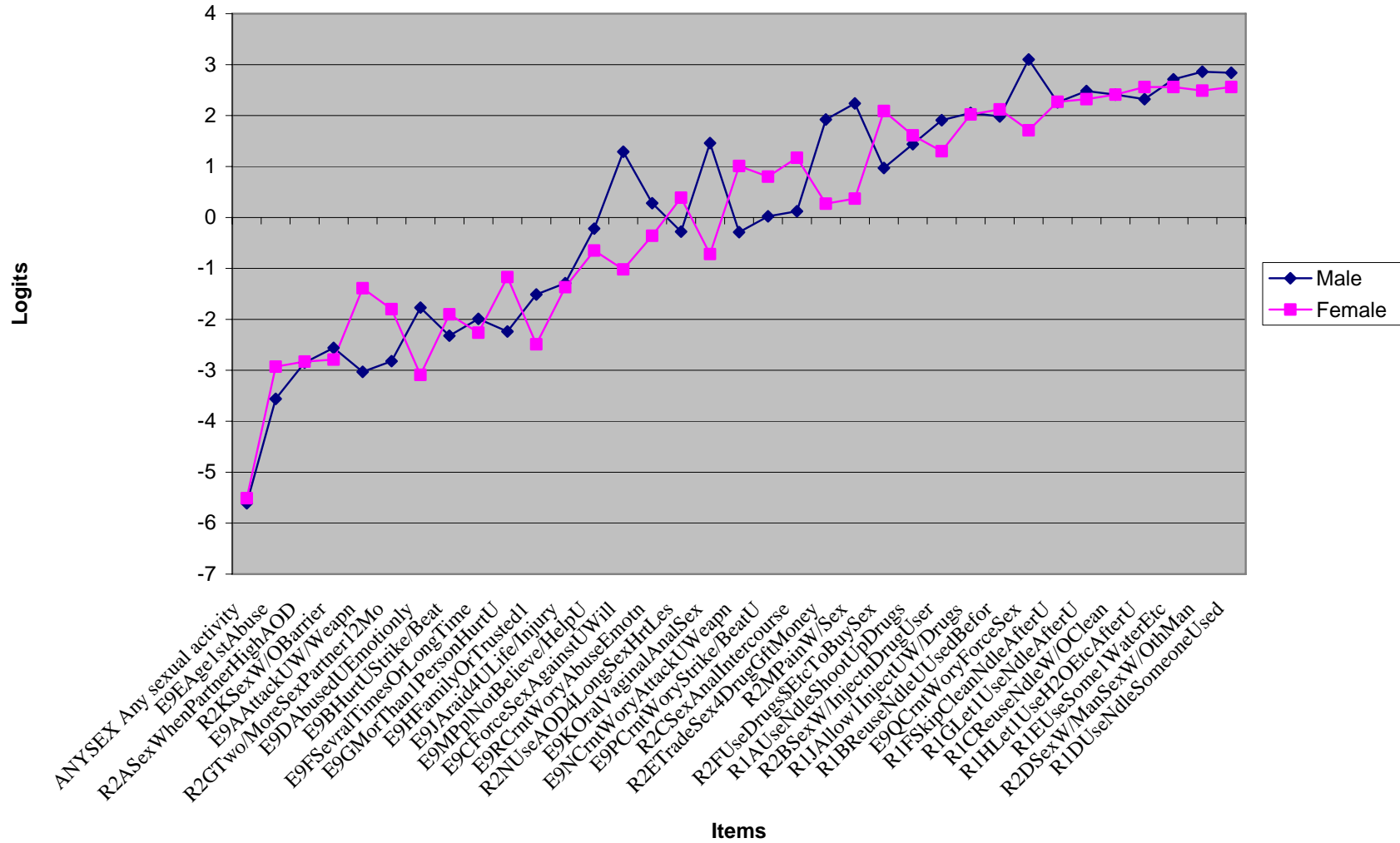


Figure 3. HIV Age DIF

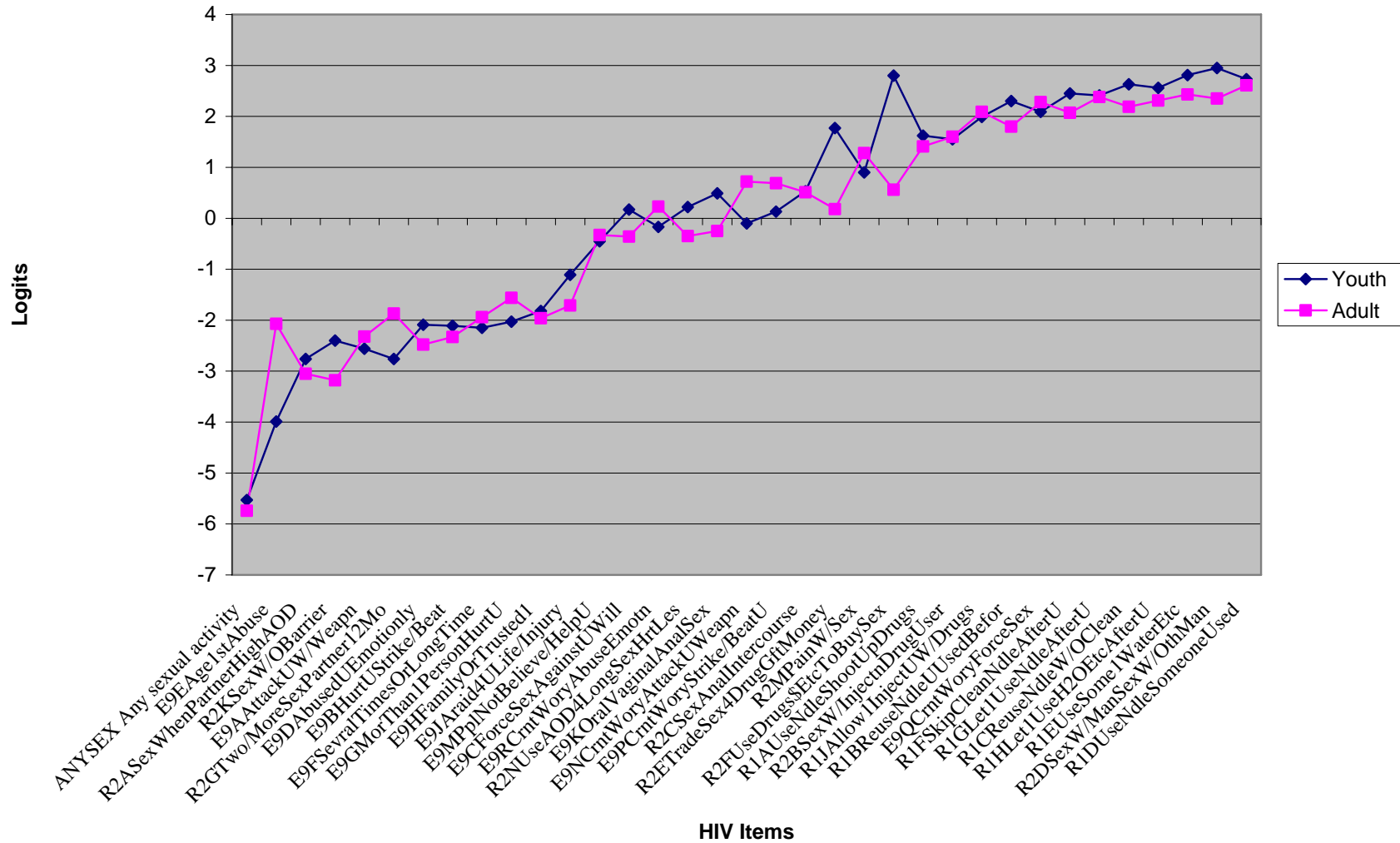


Figure 4. Race DIF

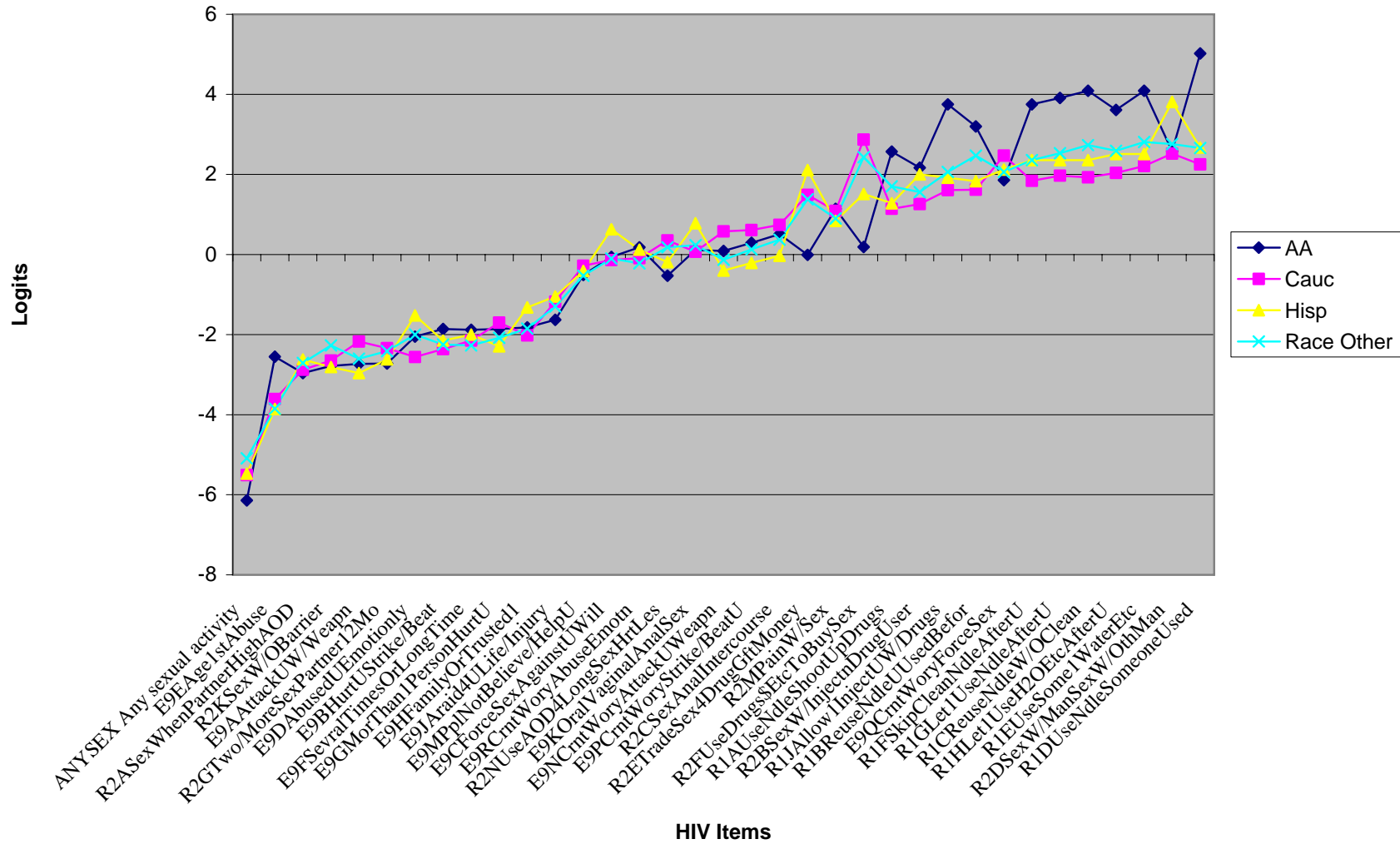
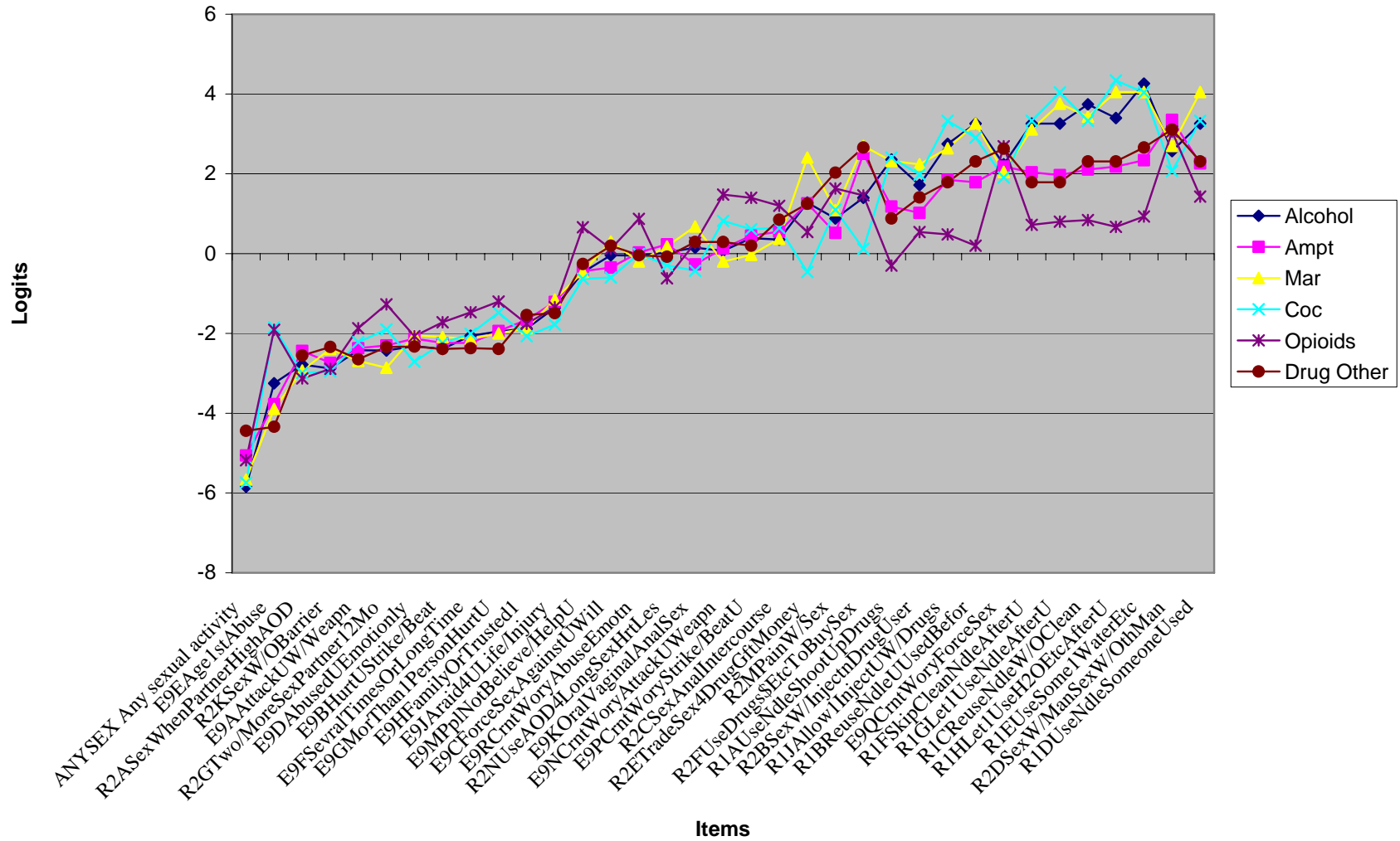


Figure 5. HIV Primary Drug DIF



**Table 13. HIV Risk Item Measures by Demographic Groups (Listed in Severity Order)**

Entry	ITEMS	MEASURE	Male	Female	Youth	Adult	AA	Cauc	Hisp	Race Other	Alcohol	Amp	Mar	Coc	Opioids	Drug Other
37	ANYSEX Any sexual activity	-5.58	-5.61	-5.51	-5.53	-5.74	-6.14	-5.51	-5.47	-5.09	-5.84	-5.06	-5.66	-5.74	-5.18	-4.44
5	E9EAge1stAbuse	-3.36	-3.56	-2.93	-3.99	-2.07	-2.55	-3.61	-3.86	-3.85	-3.25	-3.77	-3.89	-1.85	-1.91	-4.34
25	R2ASexWhenPartnerHighAOD	-2.84	-2.85	-2.83	-2.76	-3.05	-2.96	-2.88	-2.62	-2.71	-2.78	-2.44	-2.93	-2.99	-3.13	-2.55
34	R2KSexW/OBarrier	-2.63	-2.56	-2.79	-2.4	-3.18	-2.78	-2.65	-2.81	-2.27	-2.88	-2.74	-2.39	-2.95	-2.89	-2.34
1	E9AAttackUW/Weapn	-2.49	-3.03	-1.39	-2.56	-2.32	-2.73	-2.17	-2.96	-2.6	-2.42	-2.37	-2.69	-2.22	-1.87	-2.65
31	R2GTwo/MoreSexPartner12Mo	-2.48	-2.82	-1.8	-2.76	-1.87	-2.72	-2.34	-2.62	-2.41	-2.43	-2.3	-2.86	-1.9	-1.27	-2.34
4	E9DAbusedUEmotionly	-2.21	-1.77	-3.09	-2.09	-2.48	-2.05	-2.56	-1.52	-2	-2.32	-2.13	-2.05	-2.71	-2.07	-2.33
2	E9BHurtUStrike/Beat	-2.18	-2.32	-1.9	-2.11	-2.33	-1.86	-2.36	-2.15	-2.25	-2.39	-2.24	-2.09	-2.24	-1.72	-2.39
6	E9FSevralTimesOrLongTime	-2.08	-1.99	-2.26	-2.15	-1.94	-1.88	-2.15	-2	-2.27	-2.06	-2.24	-2.14	-2.01	-1.47	-2.37
7	E9GMorThan1PersonHurtU	-1.88	-2.24	-1.17	-2.03	-1.56	-1.86	-1.7	-2.29	-2.08	-1.95	-1.94	-2	-1.48	-1.2	-2.39
8	E9HFamilyOrTrusted1	-1.86	-1.51	-2.49	-1.82	-1.96	-1.82	-2.02	-1.32	-1.85	-1.88	-1.66	-1.87	-2.07	-1.76	-1.54
9	E9JArad4ULife/Injury	-1.32	-1.29	-1.37	-1.11	-1.71	-1.63	-1.18	-1.05	-1.3	-1.36	-1.21	-1.15	-1.78	-1.34	-1.49
11	E9MPpINotBelieve/HelpU	-0.41	-0.22	-0.65	-0.45	-0.33	-0.51	-0.28	-0.41	-0.53	-0.47	-0.44	-0.46	-0.63	0.66	-0.26
3	E9CForceSexAgainstUWill	-0.05	1.29	-1.02	0.17	-0.36	-0.06	-0.14	0.63	-0.1	-0.04	-0.35	0.3	-0.6	0.12	0.2
15	E9RCrntWoryAbuseEmotn	-0.03	0.28	-0.36	-0.17	0.23	0.18	-0.09	0.12	-0.22	-0.05	0.03	-0.2	-0.01	0.87	-0.04
36	R2NUseAOD4LongSexHrtLes	-0.01	-0.28	0.39	0.22	-0.35	-0.53	0.35	-0.19	0.18	0.03	0.23	0.18	-0.29	-0.62	-0.08
10	E9KOralVaginalAnalSex	0.16	1.46	-0.72	0.49	-0.25	0.11	0.07	0.78	0.23	0.15	-0.27	0.68	-0.43	0.27	0.29
12	E9NCrntWoryAttackUWeapn	0.19	-0.29	1.01	-0.1	0.72	0.09	0.58	-0.4	-0.13	0.07	0.13	-0.2	0.82	1.48	0.29
13	E9PCrntWoryStrike/BeatU	0.34	0.02	0.8	0.13	0.69	0.29	0.61	-0.21	0.13	0.39	0.45	-0.03	0.61	1.4	0.2
27	R2CSEXAnallIntercourse	0.52	0.12	1.17	0.53	0.51	0.51	0.74	-0.03	0.37	0.35	0.55	0.37	0.65	1.2	0.85
29	R2ETradeSex4DrugGftMoney	0.91	1.92	0.27	1.77	0.18	-0.01	1.49	2.11	1.38	1.28	1.26	2.41	-0.46	0.54	1.25
35	R2MPainW/Sex	1.05	2.24	0.37	0.9	1.28	1.14	1.09	0.84	0.9	0.88	0.52	1.09	1.1	1.63	2.03
30	R2FUseDrugs\$EtcToBuySex	1.42	0.97	2.09	2.8	0.56	0.19	2.87	1.51	2.43	1.4	2.5	2.7	0.12	1.46	2.66
16	R1AUseNdleShootUpDrugs	1.53	1.44	1.61	1.62	1.41	2.57	1.14	1.28	1.7	2.36	1.18	2.32	2.41	-0.3	0.88
26	R2BSEXW/InjectnDrugUser	1.57	1.91	1.3	1.55	1.6	2.17	1.26	2	1.56	1.72	1.02	2.24	1.98	0.54	1.41
24	R1JAllow1InjectUW/Drugs	2.04	2.05	2.02	1.99	2.09	3.75	1.61	1.92	2.06	2.75	1.85	2.64	3.33	0.48	1.79
17	R1BReuseNdleUUsedBefor	2.06	1.98	2.12	2.3	1.8	3.2	1.62	1.83	2.47	3.26	1.79	3.26	2.91	0.2	2.31
14	E9QCrntWoryForceSex	2.18	3.1	1.71	2.09	2.28	1.86	2.47	2.12	2.06	2.24	2.18	2.05	1.91	2.69	2.63
21	R1FSkipCleanNdleAfterU	2.27	2.26	2.27	2.45	2.07	3.75	1.84	2.36	2.36	3.26	2.04	3.11	3.33	0.72	1.79
22	R1GLet1UseNdleAfterU	2.4	2.48	2.32	2.41	2.38	3.91	1.97	2.36	2.53	3.26	1.97	3.76	4.04	0.8	1.79
18	R1CReuseNdleW/OClean	2.41	2.41	2.41	2.63	2.19	4.09	1.93	2.36	2.73	3.74	2.11	3.43	3.33	0.84	2.31

Entry	ITEMS	MEASURE	Male	Female	Youth	Adult	AA	Cauc	Hisp	Race Other	Alcohol	Amp	Mar	Coc	Opioids	Drug Other
23	R1HLet1UseH2OEtcAfterU	2.44	2.32	2.56	2.56	2.31	3.61	2.04	2.51	2.59	3.4	2.18	4.05	4.34	0.67	2.31
20	R1EUseSome1WaterEtc	2.62	2.71	2.56	2.81	2.43	4.09	2.21	2.51	2.81	4.26	2.34	4.05	4.04	0.93	2.66
28	R2DSexW/ManSexW/OthMan	2.65	2.86	2.49	2.95	2.35	2.58	2.52	3.81	2.76	2.57	3.35	2.7	2.07	3.06	3.11
19	R1DUseNdleSomeoneUsed	2.68	2.84	2.56	2.73	2.61	5.02	2.25	2.67	2.66	3.26	2.26	4.05	3.33	1.43	2.31

**Codes:** AA=African American, Cauc=Caucasian, Hisp=Hispanic, Amp=Amphetamine, Mar=Marijuana, Coc=Cocaine

## Person Fit Group Analysis

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. This information should enable us to interpret certain unusual patterns of scores more appropriately, e.g., low scorers who are actually at high risk because of their needle use. 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) so they are upside down from the typical Rasch charts where more rare is higher.

Figure 6 displays the numbers and percentages of persons in each fit group. In Figures 7-10 below, we present the four possible person fit patterns using Rasch person fit statistics (Wright & Stone, 1979) where  $\leq 1.33$  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 or poor fitting patterns (HI). 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 nearly 80% of the persons 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 expected and actual.

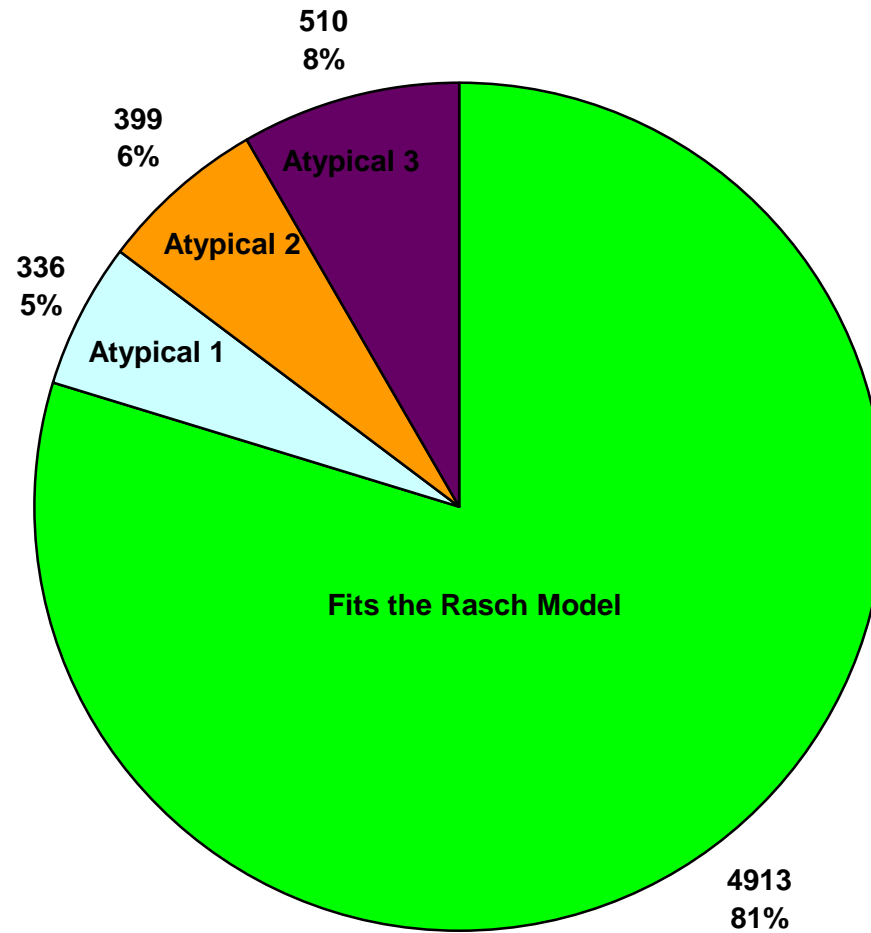
In Figure 8, the L/M infit and HI outfit group (5.4%) is called Atypical Type 1, where the overall score may underestimate severity since these tend to be people who are having risky sex (higher risk) but are unexpectedly low on victimization (which tends to be lower risk).

In Figure 9, the HI infit and L/M outfit group (6.5%) is called Atypical Type 2 and tends to be valid low scorers who endorse a few victimization items. These measures may validly be interpreted as they are.

In Figure 10, the HI infit and HI outfit group (8.3%) called Atypical Type 3 tend to be valid high scorers who also have an overall measure that may underestimate severity slightly and tend to be needle users with risky sex and victimization. These persons misfit because they endorse most of the higher risk items but do not endorse some of the lower risk items.

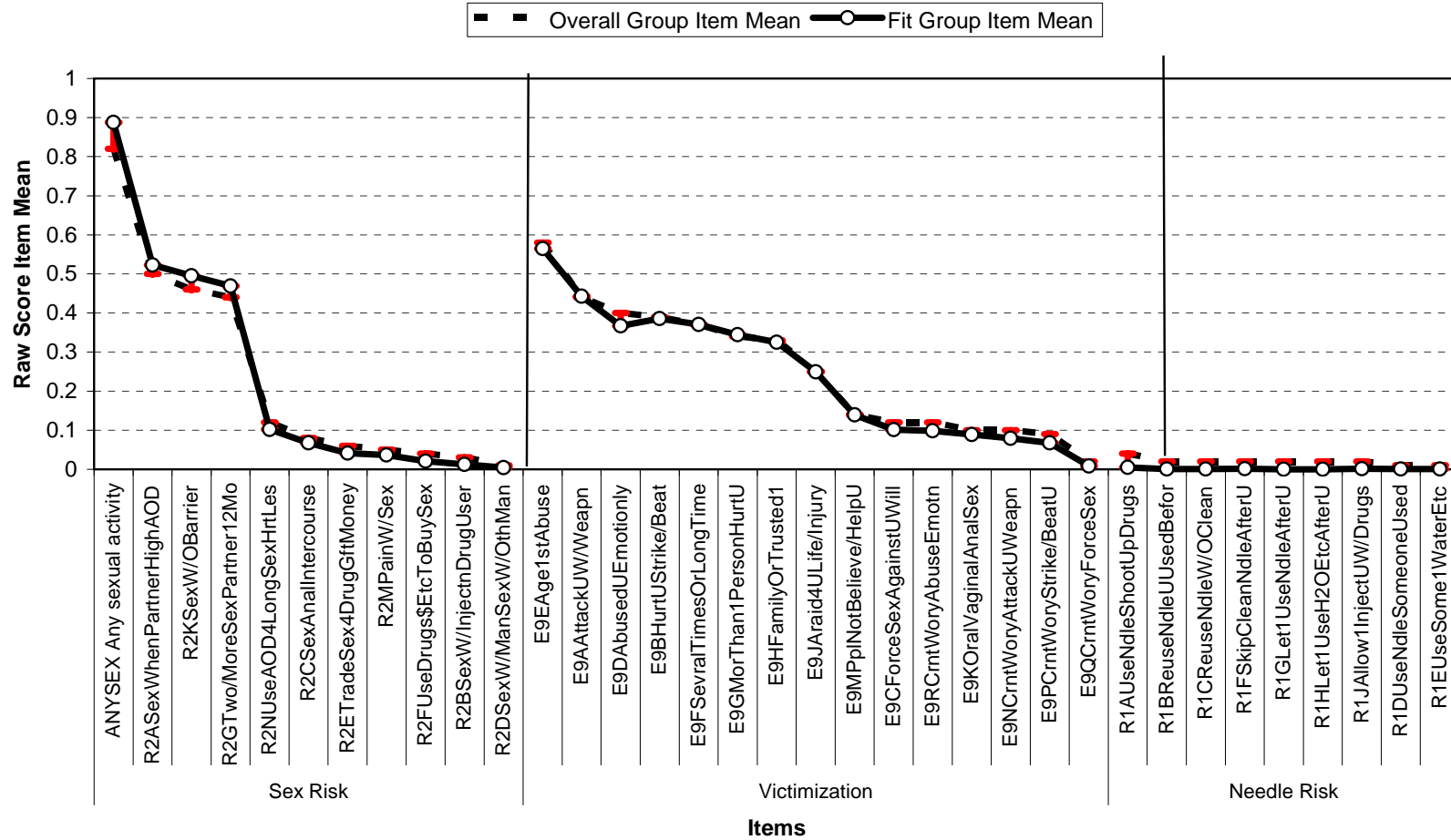
In summary, the pattern that is most deceptive is Atypical Type 1 since the person measures will tend to be substantially lower than they should be based on the person's risky sexual behavior. Atypical Type 3 will also tend to be somewhat deceptive since, even though they will tend to have high measures, their actual measures should be even higher because they tend to engage in the most risky behaviors.

Figure 6. HIV Fit Groups

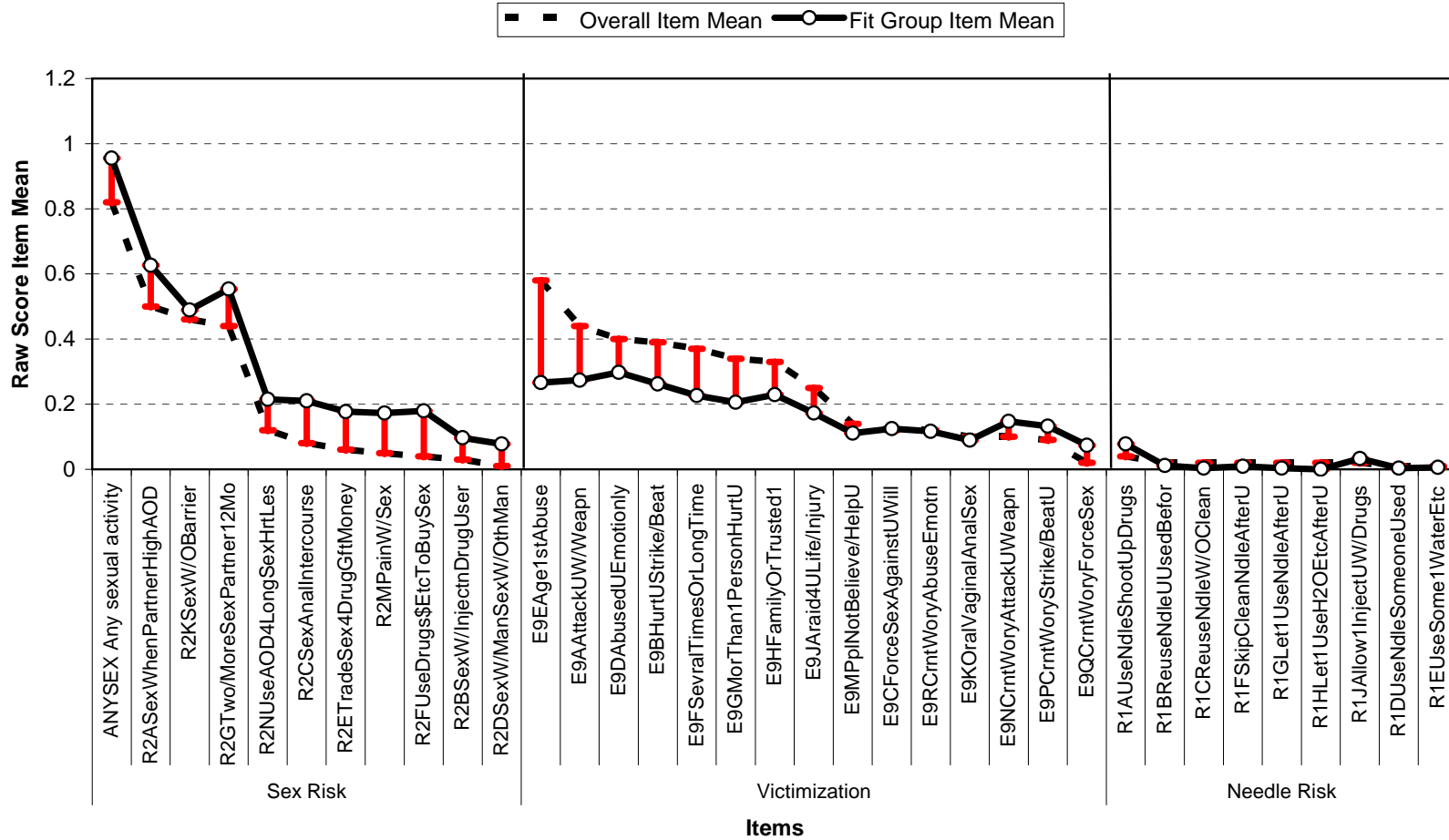


■ Fits Rasch Model (Low/Mod Infit and Low/Mod Outfit) □ Atypical 1 (Low/Mod infit and High Outfit)  
■ Atypical 2 (High Infit and Low/Mod Outfit) ■ Atypical 3 (High Infit/High Outfit)

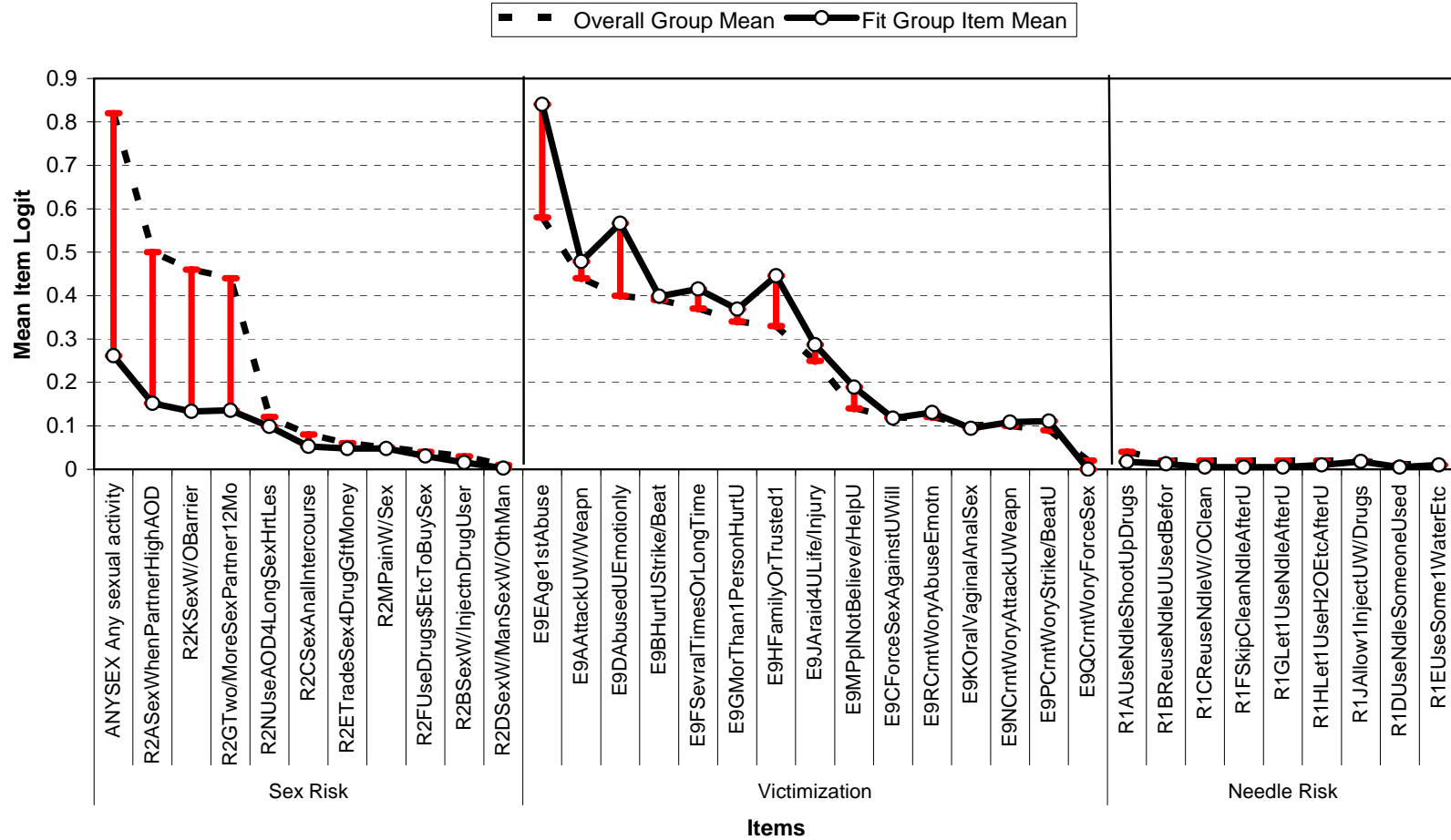
**Figure 7. HIV Overall Group vs. Fit Group Means:  
 Fits Rasch Model (n=4913; 80%)  
 (Low/Mod on Infit and Low/Mod on Outfit )**



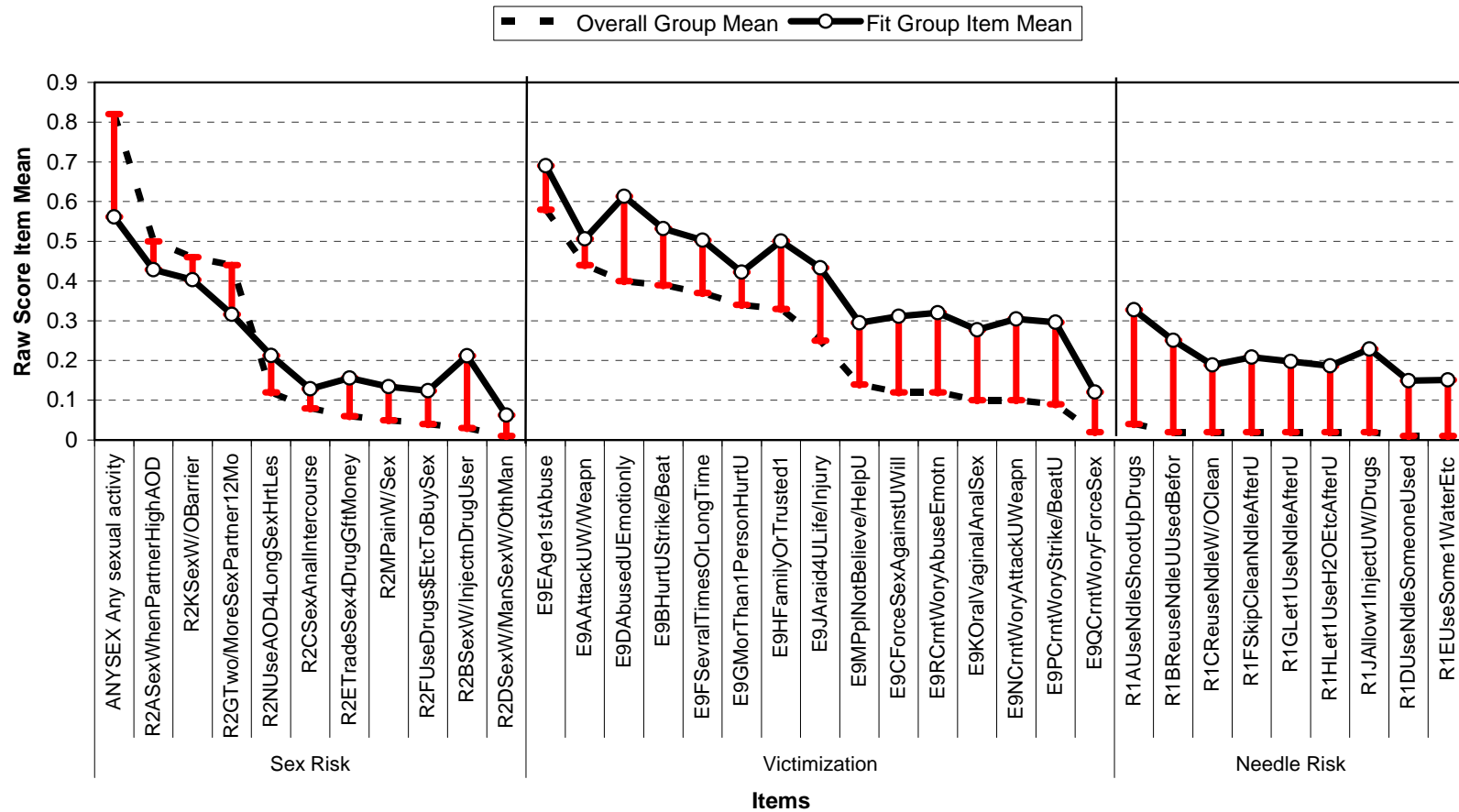
**Figure 8. HIV Overall Group vs. Fit Group Means:  
Atypical Type 1 With HIGHER Than Expected Sex Risk and LOWER Victimization (n=336; 5%)  
(Low/Mod Infit and High Outfit)**



**Figure 9. HIV Overall Group vs. Fit Group Means:  
 Atypical Type 2 With LOWER Than Expected Sex Risk and HIGHER Victimization (n=399; 6%)  
 ( High Infit and Low/Mod Outfit)**



**Figure 10. HIV Overall Group vs. Fit Group Means:  
 Atypical Type 3 with LOWER than Expected on Less Severe Sex Risk and HIGHER than  
 Expected on More Severe Sex Risk, Victimization, and Needle Use (n=510; 8%)  
 (High Infit and High on Outfit)**



### Recommended Actions and Deliberations

The *HIV Risk* scale is useful in assessing its target construct. However, there are several actions and deliberations that might improve it.

- Improve targeting of the *HIV Risk* items by developing lower risk items and by filling in gaps in the item hierarchy.
- There was clinically significant DIF (> half a standard deviation) in several analyses. While such DIF usually cancels out in the overall scale, it should be taken into account when interpreting items and for tailoring interventions to these groups. In some cases, measures may need to be adjusted when the goal is to compare these groups.
- Further analyses of DIF and its consequences for measurement of subgroups should be considered.
- Persons having both infit and outfit that is low or moderate are regarded as fitting well from a clinical perspective, so their measures are interpretable in the normal way, i.e., low measures indicate less and high indicate more of the construct.
- The three atypical groups, Atypical Types 1, 2 and 3, tend to have scores that underestimate HIV risk. We recommend flagging these groups for clinicians in the evaluation/validity concerns section of the GAIN Recommendation and Referral Summary (GRRS) as:
  - Atypical Type 1 response patterns on the HIV Scale, (Relative to total score, higher than expected engagement in high risk sex, lower than expected on the lower risk victimization items)
  - Atypical 2 response pattern on HIV Scale (Relative to total score, lower than expected on the lower sex risk items, higher than expected on a few victimization items)
  - Atypical 3 response pattern on HIV Scale (Relative to total score, lower than expected on the lower sex risk items, higher than expected on victimization and needle use)
- More work on construct validity would be helpful to understand these fit groups better and to ensure proper interpretation of measures.

March 20, 2009

## References

Bond, T.G. & Fox, C.M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences*. (2<sup>nd</sup> 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.

Conrad, K.J., Dennis, M.L., Bezruczko, N., Funk, R., Riley, B. (2007). Substance Use Disorder Symptoms: Evidence of Differential Item Functioning by Age. Journal of Applied Measurement, 8(4), 373-387.

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.

Dennis, M. L., Titus, J. C., White, M. K., Unsicker, J. I., & Hodgkins, D. (2003). Global Appraisal of Individual Needs (GAIN): Administration guide for the GAIN and related measures. Bloomington, IL: Chestnut Health Systems. Retrieved online on 10/1/08 from <http://www.chestnut.org/LI/gain/index.html>

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. (2007). *Winsteps Rasch Measurement* (Version 3.64.2). [www.winsteps.com](http://www.winsteps.com). Author.

Linacre J.M., Wright B.D. (1994) Chi-Square Fit Statistics. *Rasch Measurement Transactions* 8:2 p. 350.

Norman, G.R., Sloan, J.A., & Wywich, 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 Paedogogiske 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; 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.

030209