Validation of the Crime and Violence Scale (CVS) 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

Ya-Fen Chan
The University of Hong Kong

Rod Funk
Chestnut Health Systems

Abstract

Purpose. The purpose of this report is to provide a brief psychometric analysis of the *Crime and Violence Scale (CVS)* using the Rasch measurement model. The *CVS* consists of 31 yes/no items related to increasingly violent strategies used for resolving interpersonal conflict in the past year and the types of drug related, property, and interpersonal crimes the respondent has committed.
Methods. Data were analyzed on 7,435 persons who presented for substance problem screening. 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.
Results. Both items and scales form the theoretically expected hierarchies with a Rasch person internal consistency reliability of .82 and an item reliability of 1.00. Cronbach's alpha is .90. Of the 31 items in the *CVS*, significant DIF (i.e., > .5 SD = .94 logits) occurred in 4 items for males vs. females, 5 items for youth vs. adults, 2 items for race when using Caucasians as the reference group, and 2 items for primary substances when using alcohol as the referent. In terms of person fit, 79% 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. The L/M infit and HI outfit group (7%) is called Atypical Type 1, where the person measure may underestimate severity slightly since these tend to be people who report serious crimes (higher criminality) but are unexpectedly low or on target on only two items, i.e., *E8A Discussed It Calmly and Settled It* and *E8B Left Room or Area Rather than Argue*. The HI infit and L/M outfit group (2%) is called Atypical Type 2 and tends to be persons with moderate measures whose actual item means are unexpectedly lower on the less serious *Conflict Tactic Scale* items and unexpectedly higher on property crimes, fighting/hitting, and drug crimes. The HI infit and HI outfit group (12%) called Atypical Type 3 tend to be persons who are high in criminality but have an overall measure that may underestimate severity. They tend to be persons who endorse the more serious property, interpersonal, and drug crimes and not endorse the less serious *Conflict Tactic Scale* items. These persons misfit the Rasch model because they tend to endorse most of the higher seriousness crimes but do not tend to endorse some of the lower risk crimes as much as expected by the Rasch model.

Conclusion. The *CVS* performs well as a measure of the construct of "crime and violence." However, two items (*Discussed It Calmly and Settled the Disagreement* and *Left the Room or Area Rather than Argue*) misfit the CVS both psychometrically and logically having seemingly little to do with the construct of crime and violence. Two other items, *Prostitution* and *Forgery/BadChecks*, merit consideration. These items are much easier for females and adults to endorse. However, because the sample is predominantly young males who do not tend to endorse *Prostitution* and *Forgery/BadChecks*, these items appear to be much more severe/more rare crimes than they should be. More work on construct validity would be helpful. Regarding person fit, the two groups with the most misleading measures are Atypical Types 2 and 3. Both of these groups tend to have measures that underestimate the seriousness of their criminality. Therefore, clinicians should be alerted that persons in these fit groups will tend to have lower measures than would be appropriate based on their tendency to endorse more serious crimes.

Citation. Conrad, K. J., Conrad, K. M., Dennis, M. L., Riley, B.B., Chan, Y. F., & Funk, R. (2009). Validation of the Crime and Violence Scale (CVS) 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 CVS Rasch Report.pdf

Purpose of this Report

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

Background

The *CVS* is a scale of the *General Individual Severity Scale* (GISS). The GISS in turn is part of the larger *Global Appraisal of Individual Needs* (GAIN) which is a standardized biopsychosocial instrument that integrates research and clinical assessment for people presenting to substance abuse treatment or other behavioral health treatment (Dennis, Chan, & Funk, 2006).

The *CVS* is a count of increasingly violent strategies used by the respondents for resolving interpersonal conflict in the past year and of the types of drug related, property, and interpersonal crimes the respondent has committed. It is based on the *Conflict Tactic Scale* introduced in the *Family Violence Survey* and lay versions of the Federal Bureau of Investigations uniform crime report categories introduced in the 1995 *National Household Survey on Drug Abuse* and predicts future crime and violence (Dennis et al., 2006).

CVS Subscales and Items

The *CVS* consists of four subscales with a total of 31 items. Its subscales are the: *General Conflict Tactic Scale* (GCTS; 12 items), *Property Crime Scale* (PCS; 7 items), *Interpersonal Crime Scale* (ICS; 7 items), and the *Drug Crime Scale* (DCS; 5 items).The item stem reads: *"During the past 12 months, have you had a disagreement in which you did the following things?"* Response format is Yes/No (coded: no=0, yes=1). The scale names, item stems, GAIN item numbers, Rasch output item location codes, and item labels are shown in the table below.

Table 1. Scale and Item Information

Subscale Name and Item Stem	*GAIN* Item Number	Rasch Output Item Number	Item Label
General Conflict Tactic Scale			
1. Discussed it calmly and settled the disagreement?	E8A	94	DiscussedItCalmlySettled it
2. Left the room or area rather than argue?	E8B	95	LeftRoomOrAreaRatherThanArgue
3. Insulted, swore, or cursed at someone?	E8C	96	InsultedOrSworeAtSomeone
4. Threatened to hit or throw something at another person?	E8D	97	ThreatenedToHitOrThrowSomething
5. Actually threw something at someone?	E8E	98	ActuallyThrewSomethingAtSomeone
6. Pushed, grabbed, or shoved someone?	E8F	99	PushedGrabbedOrShovedSomeone
7. Slapped another person?	E8G	100	SlappedAnotherPerson
8. Kicked, bit, or hit someone?	E8H	101	KickedBitOrHitSomeone
9. Hit or tried to hit anyone with something (an object)?	E8J	102	HitOrTriedAnyoneWithSomething
10. Beat up someone?	E8K	103	BeatUpSomeone
11. Threatened anyone with knife or gun?	E8M	104	ThreatenedAnyoneGunOrKnife
12. Actually used a knife or gun on someone?	E8N	105	ActuallyUsedGunOrKnifeOnSomeone
Property Crime Scale			
13. Purposely damaged or destroyed property that did not belong to you?	L3A1D	106	PropertyDamage
14. Bought, received, possessed, or stolen goods?	L3A2D	107	PossessStolenGoods
15. Passed bad checks, forged, or altered a prescription, or took money from an employee?	L3A3D	108	Forgery/badChecks
16. Taken something from a store without paying for it?	L3A4D	109	Theft (store)
17. Other than from a store, taken money or property that didn't belong to you?	L3A5D	110	Theft (other)
18. Broken into a house or building to steal something or just to look around?	L3A6D	111	BreakAndEnter
19. Taken a car that didn't belong to you?	L3A7D	112	VehicleTheft
Interpersonal Crime Scale			
20. Used a weapon, force, or strong-arm methods to get money or things from a person?	L3A8D	113	ArmedTheft(money)
21. Hit someone or gotten into a physical fight?	L3A9D	114	Fighting/hitting
22. Hurt someone badly enough they needed bandages or a doctor?	L3A10D	115	Hurt other=need medical attn
23. Used a knife or gun or some	L3A11D	116	Armed theft (other)

Subscale Name and Item Stem	*GAIN* Item Number	Rasch Output Item Number	Item Label
other thing, like a club, to get something from a person?			
24. Made someone have sex with you by force when they did not want to have sex?	L3A12	117	Rape
25. Been involved in the death or murder of another person (including accidents)?	L3A13D	118	Homicide (even accident)
26. Intentionally set a building, car, or other property on fire?	L3A14D	119	Arson
Drug Crime Scale			
27. Driven a vehicle while under the influence of alcohol or illegal drugs?	L3A15D	120	DUI
28. Sold, distributed, or helped to make illegal drugs?	L3A16D	121	Sell/make drugs
29. Traded sex for food, drugs, or money?	L3A17D	122	Prostitution
30. Been a member of a gang?	L3A18D	123	Gang member
31. Gambled illegally?	L3A19D	124	Illegal gambling

Data Source

Data on the 7435 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 *CVS* 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 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

The data for this analysis come from 7435 respondents who completed the **CVS**. The respondents were being screened for substance use disorders. In the previous year, 86% had substance use disorders, 51% had internalizing disorders (e.g., somatic, depression, anxiety, trauma, suicide), 59% had externalizing disorders (i.e., attention-deficit or hyperactivity disorders, and conduct disorders), and 59% had problems with crime or violence. Approximately, 42% were entering residential treatment and 66% were involved in the criminal justice system.

As shown in the following table, the sample was predominately under 18 years of age (73%) and male (67%). Almost half were Caucasian (45%), a quarter were African American (26%), and the remainder Hispanic or mixed race. Of the top five primary drugs reported, marijuana was reported by 49% of the sample. The primary drug least often reported was opioids at 5%. Other primary drugs reported included amphetamines (11%), cocaine (11%), and alcohol (20.5%). Almost 3% percent of the sample reported other drugs.

Table 2. Demographic Characteristics of the Sample (N=7435[a])

	Percent	Number
Age, Mean (sd) 19.9 (8.9)		
< 18 years	72.5	5388
≥18 years	27.5	2047
Gender		
Male	67.1	4992
Female	32.7	2437
Race		
African American	25.7	1913
Caucasian	45.2	3360
Hispanic	10.8	806
Mixed/other	17.7	1314
Drug, primary, most severe		
Alcohol	20.6	1527
Amphetamines	11.0	820
Marijuana	49.1	3654
Cocaine	10.9	808
Opiates	5.3	393
Other drug	2.9	214

[a] Numbers may not add up to 100% due to missing values

Table 3. Person and Item Reliability

SUMMARY OF 6022 MEASURED (NON-EXTREME) **PERSONS**

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	8.6	30.1	-1.70	.58	.97	-.1	1.14	.5
S.D.	5.6	1.8	1.53	.15	.39	1.2	1.45	.8
MAX.	29.0	31.0	4.02	1.41	2.60	5.7	9.90	6.2
MIN.	1.0	3.0	-4.73	.45	.39	-2.9	.09	-1.3
REAL RMSE	.64	ADJ.SD	1.38	SEPARATION	2.16	PERSON RELIABILITY		.82
MODEL RMSE	.60	ADJ.SD	1.40	SEPARATION	2.36	PERSON RELIABILITY		.85
S.E. OF PERSON MEAN = .02								

MAXIMUM EXTREME SCORE: 1 PERSONS
MINIMUM EXTREME SCORE: 1367 PERSONS
LACKING RESPONSES: 45 PERSONS
VALID RESPONSES: 97.2%

SUMMARY OF 7390 MEASURED (EXTREME AND NON-EXTREME) **PERSONS**

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	7.0	30.1	-2.50	.82				
S.D.	6.1	2.0	2.18	.52				
MAX.	29.0	31.0	4.02	2.18				
MIN.	.0	1.0	-6.07	.45				
REAL RMSE	.99	ADJ.SD	1.94	SEPARATION	1.95	PERSON RELIABILITY		.79
MODEL RMSE	.97	ADJ.SD	1.95	SEPARATION	2.01	PERSON RELIABILITY		.80
S.E. OF PERSON MEAN = .03								

PERSON RAW SCORE-TO-MEASURE CORRELATION = .94 (approximate due to missing data)
CRONBACH ALPHA (KR-20) PERSON RAW SCORE RELIABILITY = .90 (approximate due to missing data)

- Good person reliability (.82)
- Including extreme scores, the reliability is .79
- Two separation levels split the persons into 3 groups on the Rasch ruler (see Figure 1)
- Cronbach's alpha = .90
- Cronbach's alpha is higher (.90) because it estimates extreme scores as measured perfectly, i.e., with no error

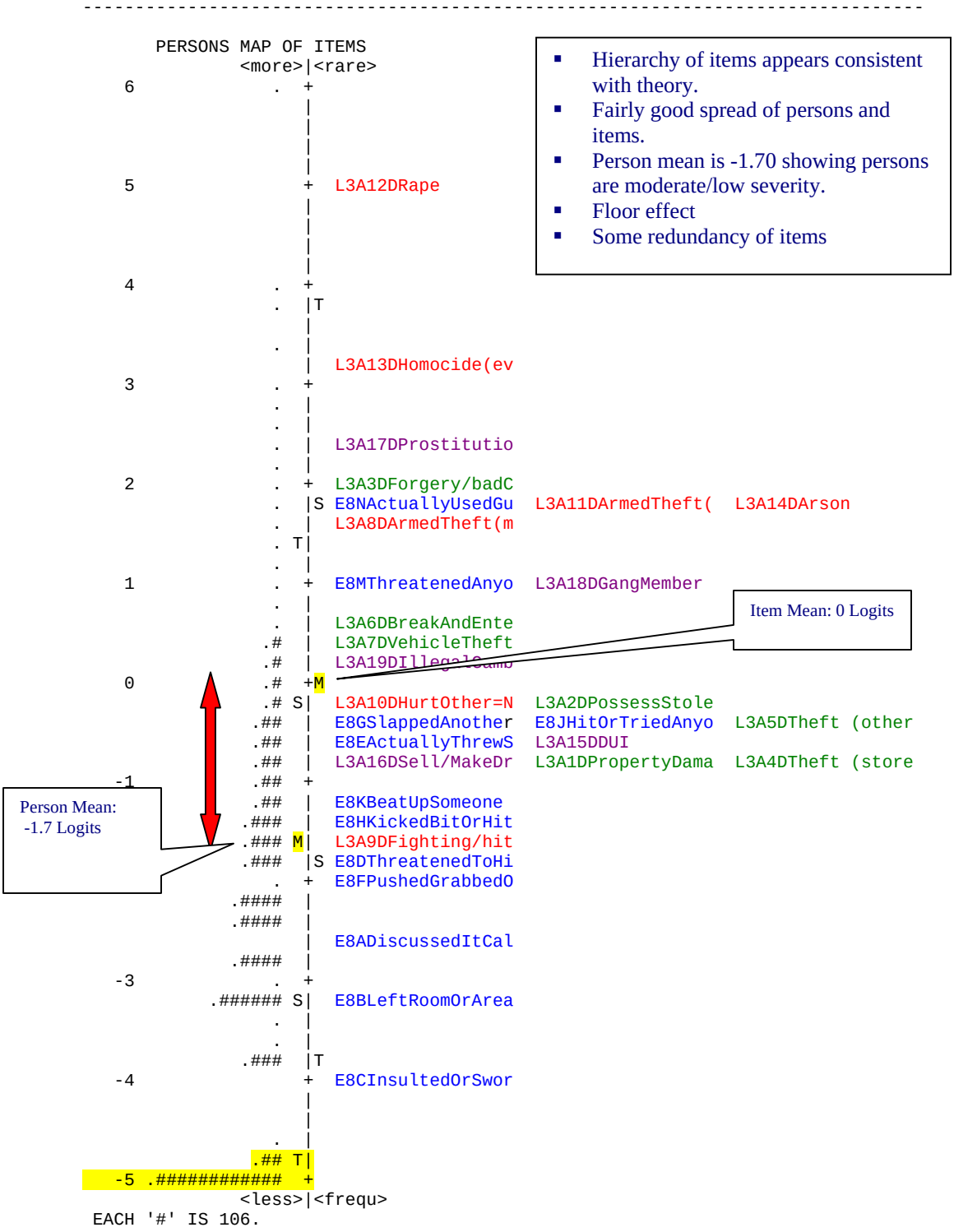
SUMMARY OF 31 MEASURED (NON-EXTREME) **ITEMS**

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	1666.7	5855.1	.00	.05	.99	-1.7	1.25	-.1
S.D.	1278.9	496.4	1.87	.03	.15	5.3	.82	6.1
MAX.	5036.0	6013.0	5.07	.21	1.52	9.9	4.18	9.9
MIN.	24.0	3963.0	-3.97	.03	.79	-9.9	.52	-9.8
REAL RMSE	.06	ADJ.SD	1.87	SEPARATION	31.86	ITEM RELIABILITY		1.00
MODEL RMSE	.06	ADJ.SD	1.87	SEPARATION	32.51	ITEM RELIABILITY		1.00
S.E. OF ITEM MEAN = .34								

UMEAN=.000 USCALE=1.000
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.93 (approximate due to missing data)

- Good item reliability (1.00)
- Item separation is high at 31.86 meaning items are placed reliably on the ruler.

Figure 1. Wright Item Map



- Hierarchy of items appears consistent with theory.
- Fairly good spread of persons and items.
- Person mean is -1.70 showing persons are moderate/low severity.
- Floor effect
- Some redundancy of items

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     =         56.2 100.0%         100.0%
  Raw variance explained by measures   =         25.2  44.8%          43.9%
    Raw variance explained by persons  =         12.0  21.3%          20.9%
    Raw Variance explained by items    =         13.2  23.5%          23.1%
  Raw unexplained variance (total)     =         31.0  55.2% 100.0%   56.1%
    Unexplned variance in 1st contrast =          3.4   6.1%  11.0%
```

STANDARDIZED RESIDUAL LOADINGS FOR ITEMS (SORTED BY LOADING)

CON-TRAST	LOADING	INFIT MEASURE	INFIT MNSQ	OUTFIT MNSQ	ENTRY NUMBER	ITEM
1	.57	-2.01	.83	.75	A 99	E8FPushedGrabbedOrShovedSomeon
1	.51	-1.40	.87	.80	B 101	E8HKickedBitOrHitSomeone
1	.48	-.67	.97	.84	C 98	E8EActuallyThrewSomethingAtSom
1	.48	-1.86	.90	.88	D 97	E8DThreatenedToHitOrThrowSomet
1	.43	-.40	1.09	.95	E 100	E8GSlappedAnotherPerson
1	.41	-.32	.91	.74	F 102	E8JHitOrTriedAnyoneWithSomethi
1	.41	-3.97	.95	1.19	G 96	E8CInsultedOrSworeAtSomeone
1	.39	-1.17	.87	.76	H 103	E8KBeatUpSomeone
1	.20	.98	.92	.74	I 104	E8MThreatenedAnyoneGunOrKnife
1	.15	-3.19	1.32	3.37	J 95	E8BLeftRoomOrAreaRatherThanArg
1	.13	1.78	.93	.87	K 105	E8NActuallyUsedGunOrKnifeOnSom
1	.11	-2.65	1.52	3.01	L 94	E8ADiscussedItCalmlySettledIt
1	-.47	-.44	.97	1.02	a 110	L3A5DTheft (other)
1	-.40	-.74	.95	1.14	b 121	L3A16DSell/MakeDrugs
1	-.40	-.21	.93	.94	c 107	L3A2DPossessStolenGoods
1	-.40	-.73	1.07	1.37	d 109	L3A4DTheft (store)
1	-.37	.51	.97	1.17	e 111	L3A6DBreakAndEnter
1	-.37	-.53	1.11	1.55	f 120	L3A15DDUI
1	-.35	-.71	.91	1.00	g 106	L3A1DPropertyDamage
1	-.34	.40	1.00	1.24	h 112	L3A7DVehicleTheft
1	-.25	1.53	.84	.59	i 113	L3A8DArmedTheft(money)
1	-.24	.18	1.03	1.18	j 124	L3A19DIllegalGambling
1	-.22	1.86	.82	.52	k 116	L3A11DArmedTheft(other)
1	-.22	-1.52	.79	.76	l 114	L3A9DFighting/hitting
1	-.21	1.84	.98	.85	m 119	L3A14DArson
1	-.20	1.01	1.01	1.14	n 123	L3A18DGangMember
1	-.19	1.91	1.11	2.17	o 108	L3A3DForgery/badChecks
1	-.16	-.12	.85	.75	P 115	L3A10DHurtOther=NeedMedicalAtt
1	-.13	3.22	.98	.93	O 118	L3A13DHomocide(even accident)
1	-.10	2.34	1.19	4.20	N 122	L3A17DProstitution
1	-.04	5.07	1.04	1.44	M 117	L3A12DRape

- To judge the strength of the measurement dimension, we used the following internal guidelines for variance explained by the measure: ≥ 40% is considered a strong measurement dimension, ≥30% is considered a moderate measurement dimension, and ≥ 20% is considered a minimal dimension. The 20% criterion is taken from Reckase (1979).
- The variance explained by the measure is 44.8%.
- The fact that only 11% of the variance is explained by the first factor of residuals supports unidimensionality
- The first factor of residuals is primarily composed of the *Conflict Tactic subscale* contrasted with the more serious crimes. We interpret this as two levels of a single construct rather than two separate constructs.
- In terms of misfit, items with high Infit mean squares (MNSQ) show a confused or random pattern that is more serious than Outfit MNSQ and reflects that these items are poor indicators of the construct.
- The items with high Outfit MNSQ are items with more unexpected responses than are consistent with the model.
- The two most misfitting items are *E8ADiscussedItCalmlySettledIt* and *E8BLeftRoomOrAreaRathrThanArg* in terms of Infit and Outfit. These are confusing items that seem to have little to do with crime and violence and could be dropped to improve the validity of the measure.
- *L3A17DProstitution* has the highest Outfit (4.18) followed by *L3A3DForgery/badChecks*, *L3A15DDUI*, *L3A12DRape* and *L3A4DTheft (store)*. These last five items have good Infit MNSQ. We interpret this to mean that the items demonstrate content validity in that they are all crimes but that some respondents are outliers who should be studied further.
- The PCA was conducted using *Winsteps v. 3.68*

Table 5. Most Unexpected Responses in Terms of Measures

```
-------------------------------------------------------------------------------

ITEM                   MEASURE  |PERSON
                                |4 447442732122 65443331 6533 775433332   66533333
                                |27472959239563491625544978435223786449995974211003
                                |99917528576287400322180899147400634869753152492520
                                |79152698605521065660053843411120428935567755754192
                              high----------------------------------------------------
  96 E8CInsultedOr    -3.97 I|..0.00...0.....................................
  95 E8BLeftRoomOr    -3.19 B|.0.0.....................................
  94 E8ADiscussedI    -2.65 C|0.....................................
 109 L3A4DTheft (s     -.73 G|...........1..1..111..1.....................
 121 L3A16DSell/Ma     -.74 M|.............1... ..1.................. .........
 106 L3A1DProperty     -.71 P|............1.............1.....................
 120 L3A15DDUI         -.52 E|............1..1.1...1..1.....................
 110 L3A5DTheft (o     -.45 O|.......................1.1.....................
 100 E8GSlappedAno     -.40 N|........1.....................................
 107 L3A2DPossessS     -.21 l|. .      .. .1..  .     .....           .........
 115 L3A10DHurtOth     -.12 e|.0......1.1..1.....................................
 112 L3A7DVehicleT      .40 H|........1.......1.....................................
 111 L3A6DBreakAnd      .50 K|........1........ .....................................
 104 E8MThreatened      .98 j|.....1.....................................
 123 L3A18DGangMem     1.02 L|0.00.......1..................11... 111.........
 105 E8NActuallyUs     1.79 k|.............................1.................
 119 L3A14DArson       1.84 o|...............1.1.........1.....................
 116 L3A11DArmedTh     1.87 b|.......................1.........................
 108 L3A3DForgery/     1.91 D|....11..1.1......1.1111..1.1......11.....1.1......
 122 L3A17DProstit     2.34 A|....1..................1.1..11.....11....1.111111
 118 L3A13DHomocid     3.22 n|... .......11111.....................................
 117 L3A12DRape        5.07 F|1...11111111.....................................
                                |----------------------------------------------low-
                                |47447442732122465443331965335775433332995665333333
                                |29472959239563491625544878437223786449753974211000
                                |99917528576287000322180899141400634869567152492522
                                |7 152698605521 65660053 4341 120428935   75575419
```

- 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 shown on the right.

- For example, Person #4297 has a high measure of 3.37 (Data are not shown here) but did not endorse the low severity item 94, “Discussed it calmly.”
- Based on our understanding and interpretation of the construct of criminality, “Discussed it calmly” does not fit logically. We interpret this as an item that does not fit the construct.
- Person 302 has a very low measure (-4.73) (Data are not shown here) but endorsed a high severity item, “Prostitution.” The large number of 1’s in the” Prostitution” row indicate that this is atypical response has occurred several times.

Table 6. Most Misfitting Response Strings in terms of Outfit Mean Squares (OUTMNSQ)

```
ITEM                 OUTMNSQ  |PERSON
                              |4 447442732122 65443331 6533 775433332   66533333
                              |27472959239563491625544978435223786449995974211003
                              |99917528576287400322180899147400634869753152492520
                              |79152698605521065660053843411120428935567755754192
                            high----------------------------------------------------
 122 L3A17DProstit        4.18 A|....1.................1.1..11.....11....1.111111
  95 E8BLeftRoomOr        3.36 B|.0.0.................................................
  94 E8ADiscussedI        3.01 C|0.................................................
 108 L3A3DForgery/        2.17 D|....11..1.1.....1.1111..1.1.....11.....1.1......
 120 L3A15DDUI            1.55 E|............1..1.1...1..1.......................
 117 L3A12DRape           1.44 F|1...11111111..................................
 109 L3A4DTheft           1.38 G|............1..1..111..1.......................
 112 L3A7DVehicleT        1.23 H|........1.......1.............................
  96 E8CInsultedOr        1.19 I|..0.00...0....................................
 111 L3A6DBreakAnd        1.17 K|........1.........  ..........................
 123 L3A18DGangMem        1.14 L|0.00........1..................11...  111.........
 121 L3A16DSell/Ma        1.14 M|..............1...  ..1.................  .........
 100 E8GSlappedAno         .95 N|.........1.....................................
 110 L3A5DTheft (o        1.02 O|.......................1.1....................
 106 L3A1DProperty        1.00 P|............1.............1....................
 119 L3A14DArson           .86 o|................1.1.........1....................
 118 L3A13DHomocid         .92 n|...  .......11111.............................
 107 L3A2DPossessS         .94 l|. .       .. .1..  .     .....           ..........
 105 E8NActuallyUs         .87 k|.................................1..............
 104 E8MThreatened         .74 j|.....1.........................................
                              |----------------------------------------------low-
                              |474474427321224654433319653357754333329956653333333
                              |29472959239563491625544878437223786449753974211000
                              |99917528576287000322180899141400634869567152492522
                              |7 152698605521 65660053 4341 120428935   7557541
```

- This table shows the most misfitting items in terms of OUTMNSQ and to the right, we see persons who unexpectedly endorsed or did not endorse an item given their overall measure.
- Prostitution has the highest outfit because this difficult item (2.34 logits) was endorsed by many people who had low overall scale measures. For example, Person #302 on the far right endorsed prostitution but had a low overall measure of -4.73. (Data are not shown here).
- The box at the left shows the seven items with high outfits. This type of finding may indicate that there are certain atypical persons with low criminality measures who commit crimes that are typically regarded as serious. Such atypical patterns are worthy of further investigation and description.

Table 7. Item Entry Order Showing Infit, Outfit, and Point-by-Measure (a.k.a. item/total) Correlations

ITEM STATISTICS: ENTRY ORDER

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
94	3918	7321	-2.65	.03	1.52	9.9	3.01	9.9	.48	.65	62.3	75.9	E8ADiscussedItCalmlySettledit	0
95	4422	7328	-3.19	.03	1.32	9.9	3.37	9.9	.56	.66	71.9	78.8	E8BLeftRoomOrAreaRatherThanArgue	0
96	5010	7332	-3.97	.04	.95	-2.4	1.19	2.8	.73	.66	87.0	85.2	E8CInsultedOrSworeAtSomeone	0
97	3122	7328	-1.86	.03	.90	-7.7	.88	-4.6	.66	.62	76.4	74.0	E8DThreatenedToHitOrThrowSomething	0
98	1933	7329	-.67	.03	.97	-2.1	.84	-5.1	.56	.54	77.8	77.6	E8EActuallyThrewSomethingAtSomeone	0
99	3285	7333	-2.01	.03	.83	-9.9	.75	-9.6	.69	.63	79.6	74.2	E8FPushedGrabbedOrShovedSomeone	0
100	1692	7328	-.40	.03	1.09	4.8	.95	-1.3	.50	.52	76.0	79.3	E8GSlappedAnotherPerson	0
101	2649	7332	-1.40	.03	.87	-9.8	.80	-8.2	.64	.59	77.6	74.5	E8HKickedBitOrHitSomeone	0
102	1622	7327	-.32	.03	.91	-5.5	.74	-7.2	.55	.51	81.6	79.9	E8JHitOrTriedAnyoneWithSomething	0
103	2412	7326	-1.17	.03	.87	-9.4	.76	-9.5	.63	.58	78.5	75.2	E8KBeatUpSomeone	0
104	750	7325	.98	.04	.92	-3.1	.74	-3.6	.42	.39	90.0	88.9	E8MThreatenedAnyoneGunOrKnife	0
105	421	7316	1.78	.06	.93	-1.7	.87	-1.3	.33	.31	93.8	93.4	E8NActuallyUsedGunOrKnifeOnSomeone	0
106	1943	7274	-.71	.03	.91	-5.5	1.00	.0	.57	.54	80.2	77.4	L3A1DPropertyDamage	0
107	960	4775	-.21	.04	.93	-3.0	.94	-1.2	.51	.49	82.7	81.1	L3A2DPossessStolenGoods	0
108	383	7290	1.91	.06	1.11	2.5	2.17	8.1	.25	.30	93.6	93.9	L3A3DForgery/badChecks	0
109	1959	7270	-.73	.03	1.07	4.3	1.37	9.9	.51	.54	76.5	77.3	L3A4DTheft (store)	0
110	1711	7277	-.44	.03	.97	-1.9	1.02	.6	.53	.52	80.2	79.0	L3A5DTheft (other)	0
111	1011	7293	.51	.04	.97	-1.1	1.17	2.6	.44	.43	86.4	85.7	L3A6DBreakAndEnter	0
112	1082	7295	.40	.04	1.00	-.2	1.24	3.8	.44	.44	85.3	85.0	L3A7DVehicleTheft	0
113	506	7289	1.53	.05	.84	-4.8	.59	-5.1	.39	.34	92.9	92.2	L3A8DArmedTheft(money)	0
114	2739	7269	-1.52	.03	.79	-9.9	.76	-9.8	.67	.60	81.6	74.3	L3A9DFighting/hitting	0
115	1437	7263	-.12	.04	.85	-8.3	.75	-6.2	.54	.49	83.7	81.3	L3A10DHurtOther=NeedMedicalAttn	0
116	396	7291	1.86	.06	.82	-4.6	.52	-5.5	.37	.31	94.2	93.7	L3A11DArmedTheft(other)	0
117	24	7298	5.07	.21	1.04	.3	1.44	3.0	.07	.09	99.6	99.6	L3A12DRape	0
118	126	7273	3.22	.10	.98	-.3	.93	-.4	.20	.19	97.9	97.9	L3A13DHomocide(even accident)	0
119	404	7286	1.84	.06	.98	-.5	.85	-1.4	.32	.31	93.6	93.6	L3A14DArson	0
120	1783	7279	-.53	.03	1.11	6.2	1.55	9.9	.48	.53	77.0	78.5	L3A15DDUI	0
121	1962	7251	-.74	.03	.95	-2.9	1.14	4.1	.55	.54	79.3	77.2	L3A16DSell/MakeDrugs	0
122	273	7293	2.34	.07	1.19	3.5	4.20	9.9	.17	.26	95.2	95.6	L3A17DProstitution	0
123	726	7277	1.01	.05	1.01	.2	1.14	1.7	.39	.39	89.0	89.2	L3A18DGangMember	0
124	768	4769	.18	.05	1.03	1.2	1.18	2.7	.44	.45	83.2	84.0	L3A19DIllegalGambling	0
MEAN	1659.0	7136.7	.00	.05	.99	-1.7	1.25	.0			84.0	83.7		
S.D.	1272.6	621.5	1.87	.03	.15	5.3	.82	6.2			8.3	7.8		

Although *E8ADiscussedItCalmlySettledIt* and *E8BLeftRoomOrAreaRatherThanArgue* have strong point-measure correlations, they are the two most misfitting items. They have the most unexpected or illogical response patterns, and they are not crimes. In comparison, items such as *L3A12DRape* and *L3A17DProstitution* have low correlations, but more logical response patterns, and they are crimes.

PERSON STATISTICS

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

```
-------------------------------------------------------------------------------
PERSON                                  MEASURE  |ITEM
                                                 |     111111 111111111111111111
                                                 |999991002009210001211021011021 1
                                                 |654974131968000275421433596828 7
                                               high-------------------------------
4297 11064103494016 1317.00 8111 42     3.37  |..0.................0......1
 799 01656104002373 2217.01 7111 79     3.33 G|.0.............. 0 ...........
4491 11414103573014 1216.01 211. 44     2.88  |0.....................0........
4715 11852104003234 1315.01 6111 47     2.56  |.0.............. . ...0...... .
7272 17121213000172 1443.01 .111 72      -.84  |0.............. . ........11.1
2988 05960208001449 1140.00 5111 29    -1.07 S|.............. . ...........1
4529 11479104001005 1217.01 2111 45    -1.07  |.............. . ...........1
4956 12357104011141 2116.01 5111 49    -1.07  |0.............. . ..1.....1..1
7256 17069212060366 1423.01 3110 72    -1.31  |.............. . 11......1..1
2965 05914205011930 2234.00 6100 29    -1.90 R|........................1..1
3370 06664209020299 1151.01 8000 33    -1.90  |0............1..1...........1
1525 03062106004131 1217.01 4110 15    -2.81 O|.............. 1 ..........11
 440 0 885103470077 1317.01 7000 44    -2.85 D|........1........1...........1.
2371 04712114130404 1316.01 5111 23    -2.85 P|..........1.....1...........1.
2682 05379114200305 1317.01 .111 26    -2.85 Q|.........1..1........1......1.
6906 16336208005476 2150.01 8101 69    -3.29  |.........1..1..............1.
 988 02015104007080 2413.00 7111 98    -3.83 M|.........1...... . .......1....
1403 02823106002276 2216.01 4110 14    -3.83 N|...........1... . ........1...
3485 06885210091743 2222.01 7001 34    -3.83  |.........1...... . ........1...
3510 06939210095300 2126.01 8101 35    -3.83  |.........1...... . ........1...
3520 06959210095889 2128.00 8011 35    -3.83  |.........1...... . ........1...
4226 1 926103470122 1315.00 7010 42    -3.82  |........ 1.......... ....1.....
4636 11681104002447 2218.01 7111 46    -3.83  |...........1... . ........1...
5105 12658106001150 1415.01 4111 51    -3.83  |.............. . 1.....1.....
 571 01162103510440 1216.01 7011 57    -3.86 F|..........1............1.....
3341 06602209020237 1143.00 8101 33    -3.86  |........................1...
3414 06753209020388 2151.01 8101 34    -3.86  |............1............1..
5893 14256111140152 1417.01 .010 58    -3.86  |...........1............1...
6794 16095208002855 2128.01 8111 67    -3.86  |............1............1..
 537 01077103500004 1116.01 7010 53    -4.71 E|........ ............1........
 956 01950104007015 1416.01 7101 95    -4.71 K|.............. . ...1........
 975 01986104007051 1417.01 7000 97    -4.71 L|.............. . ...1........
2995 05972208001573 2133.00 8111 29    -4.71 T|.............. . ... .....1..
3463 06840210069189 2130.01 8101 34    -4.71  |.............. . .........1..
3489 06895210092421 2221.01 7100 34    -4.71  |.............. . ........1...
```

- This table illustrates misfitting persons in terms of item measures.
- We highlighted Persons #4297 and #302 from the earlier examples. #4297 had a high measure of 3.37 but unexpectedly did not endorse item 94, *DiscussedItCalmlySettled it*, which we have identified as a misfitting item.

Table 9. Persons with Most Misfitting Response Strings in Terms of OUTMNSQ

```
PERSON                                  OUTMNSQ  |ITEM
                                                 |     111111 111111111111111111
                                                 |999991002009210001211021011021 1
                                                 |654974131968000275421433596828 7
                                               high----------------------------
  270 0 527103420114 1417.01 8011 27     9.90 A|.00.........................
  302 0 585103430001 2119.00 7000 30     9.90 B|.........................1..
  324 0 641103430058 1114.01 6010 32     9.90 C|00000.................1.11...1
  440 0 885103470077 1317.01 7000 44     9.90 D|........1........1...........1.
  537 01077103500004 1116.01 7010 53     9.90 E|........ ............1........
  571 01162103510440 1216.01 7011 57     9.90 F|..........1.............1.....
  799 01656104002373 2217.01 7111 79     9.90 G|.0.............. 0 ............
  801 01661104002397 1415.01 7011 80     9.90 H|.00............. . ............
  841 01743104003036 1315.01 6111 84     9.90 I|.00.0.......... . ............
  900 01853104003235 1315.01 6111 90     9.90 J|.00............. . ............
  956 01950104007015 1416.01 7101 95     9.90 K|................ . ...1........
  975 01986104007051 1417.01 7000 97     9.90 L|................ . ...1........
  988 02015104007080 2413.00 7111 98     9.90 M|.........1...... . .......1....
 1403 02823106002276 2216.01 4110 14     9.90 N|............1... . ........1...
 1525 03062106004131 1217.01 4110 15     9.90 O|................ 1 ..........11
 2371 04712114130404 1316.01 5111 23     9.90 P|..........1.....1............1.
 2682 05379114200305 1317.01 .111 26     9.90 Q|.........1..1........1......1.
 2965 05914205011930 2234.00 6100 29     9.90 R|.........................1..1
 2988 05960208001449 1140.00 5111 29     9.90 S|................ . ...........1
 2995 05972208001573 2133.00 8111 29     9.90 T|................ . ... .....1..
 3029 06025208002147 2141.00 8111 30     9.90 U|.........................1..
 3051 06069208002585 2140.01 5111 30     9.90 V|.........................1..
 3111 06181208003777 1329.01 8101 31     9.90 W|0.......1...1....1.....1..1..1.
 3124 06201208003991 2134.00 6111 31     9.90 X|.........................1..
 3155 06260208004655 2133.00 8101 31     9.90 Y|.........1................11..
 3158 06267208004723 1139.01 6111 31     9.90 Z|................1.1... .....1..
                                                 |--------------------------low-
                                                 |999991111191111111111111111111
                                                 |654971002008210001211021011021 1
                                                 |     413196 0002754214335968287
```

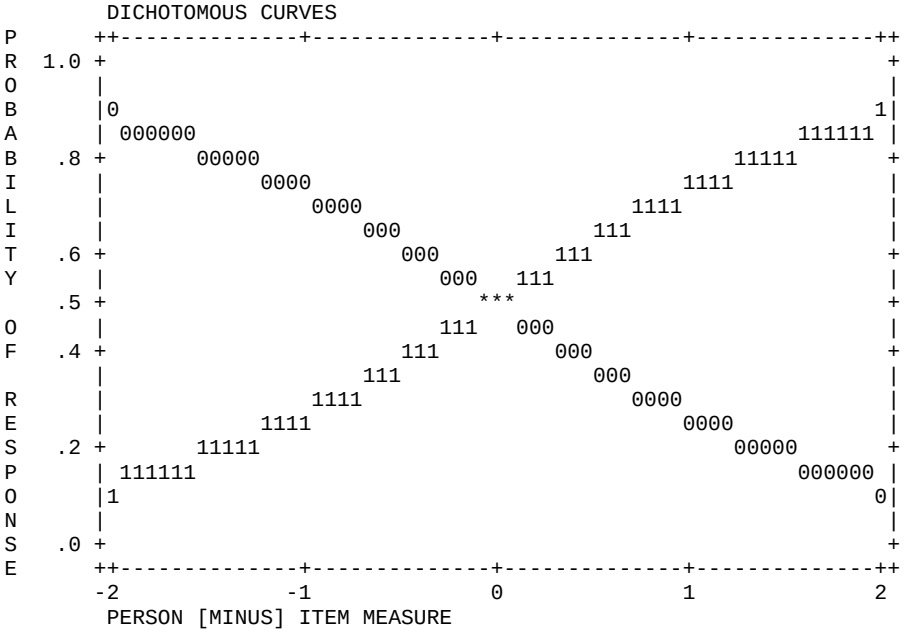
- This table shows the persons who unexpectedly endorsed an item given their overall measure. That is, we see the unexpected responses of persons with the highest outfit mean squares (MNSQ)s.
- Note, Person #302 again. #302 had a very low measure of -4.73 (Data are not shown here) but endorsed item 122, Prostitution. This resulted in a high outfit mean square of 9.90.

Table 10. Summary of Category Structure.

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	COHERENCE M->C	COHERENCE C->M	ESTIM DISCR	
0	0	129838	70	-2.62	-2.62	1.01	1.46	86%	92%		0 no
1	1	51669	28	.61	.61	.97	1.18	76%	63%	1.00	1 yes
MISSING		5175	3	-1.54							

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?

70% of all response options were answered "no", i.e., 0 was chosen as the response option.



0 = no
1 = yes

Differential Item Functioning (DIF) for Age, Gender, Race, and Primary Drug Severity for the Crime/Violence 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 analysis allows us to examine whether items have significantly different meanings for different groups. The authors 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 $\geq$ .9 logit difference for all comparisons which is approximately half a standard deviation (SD = 1.87) 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 SPS items. Table 12 contains the data that formed these graphs, and contains the information to compute differences between groups on each item. For example, to get the DIF contrast between males and females on Insulted or Swore at Someone, subtract -3.75 – (-4.44) = .69 which does not reach our chosen significance level of half a standard deviation of .9.

Gender DIF. The "Gender DIF" figure shows endorsing "*Rape*" as easier for females to endorse than it is for males is counterintuitive. It should not be easier for females to endorse this item. Given that the majority of respondents were male, the item measures (in ascending order of severity on the chart) were driven by the males. Further, males and females differentially endorsed a number of items. Thus, the females were being evaluated on a male ruler, and their scores were being weighted by the males' responses. The result is that females' scores were biased against them for a number of items. Therefore, the total measure for a female may be biased. The more serious crimes such as "*Hurting someone*", "*Armed theft*", and "*Being a gang member*" were substantially easier for males. Because the combined measures were more heavily influenced by males, this means that females were getting lower measures than they should. For females, it is substantially easier to endorse less serious offenses such as "*Slapping someone*" and "*Prostitution*." Because "*Prostitution*" is harder for males to endorse and the sample is heavily weighted by males, "*Prostitution*" is given a much higher calibration than it should be given.

However, another issue is that of the construct validity of the hierarchy overall and for males and females taken separately. For example, for females, it is one logit more difficult to endorse "Armed theft." Ordinarily, one would interpret this as meaning that "Armed theft" was more serious, i.e., more criminal and violent, for females. In other words, the usual interpretation of a higher calibration would be that a female who endorses "Armed theft" has more of the construct of Crime and Violence that is measured by the scale. However, this seems counter-intuitive. Is "Armed theft" actually regarded as more serious for women? To check the validity of this, one could examine whether women receive longer sentences for this offense than men. Does the hierarchy of actual sentences for men vs. women coincide with the Rasch hierarchy? If it does, this would be evidence of the validity of the DIF finding. If the hierarchies did not coincide, how would one interpret this DIF? As another example, is prostitution really the third most serious crime, or is this hierarchy simply being driven by the preponderance of males in the sample for whom prostitution is a very infrequent crime?

Example

In practical terms, what could this mean? Let us construct two hypothetical examples, John and Mary, where both endorse 12 out of the 31 CVS items.

- John endorsed the following items:

  8. E8H Kicked, bit, or hit someone?
  10. E8K Beat up someone?
  12. E8N Actually used a knife or gun on someone?
  18. L3A6D Broken into a house or building to steal something or just to look around?
  19. L3A7D Taken a car that didn't belong to you?
  20. L3A8D Used a weapon, force, or strong-arm methods to get money or things from a person?
  21. L3A9D Hit someone or gotten into a physical fight?
  22. L3A10D Hurt someone badly enough they needed bandages or a doctor?
  23. L3A11D Used a knife or gun or some other thing, like a club, to get something from a person?
  24. L3A12D Made someone have sex with you by force when they did not want to have sex?
  26. L3A14D Intentionally set a building, car, or other property on fire?
  28. L3A16D Sold, distributed, or helped to make illegal drugs?

- Mary endorsed the following items to get her score of 12:

  1. E8A Discussed it calmly and settled the disagreement?
  2. E8B Left the room or area rather than argue?
  3. E8C Insulted, swore, or cursed at someone?
  4. E8D Threatened to hit or throw something at another person?
  5. E8E Actually threw something at someone?
  7. E8G Slapped another person?
  16. L3A4D Taken something from a store without paying for it?
  17. L3A5D Other than from a store, taken money or property that didn't belong to you?
  27. L3A15D Driven a vehicle while under the influence of alcohol or illegal drugs?
  29. L3A17D Traded sex for food, drugs, or money?
  30. L3A18D Been a member of a gang?
  31. L3A19D Gambled illegally?

Both John and Mary got the same overall measure. The question is, even though they both endorsed the same number of items, do they both have the same level of crime and violence? The items that John endorsed were all easier for males to endorse than they were for females, i.e., highest probability that males would endorse these items. The items that Mary endorsed were all easier or equally easy for females to endorse compared to males, i.e., highest probability that females would endorse these items. Comparing John's and Mary's items, it appears clear that the severity of John's endorsements is much higher than Mary's even though they would get the same overall measure.

Age DIF. In Figure 2, however, there were not many differences, where *Prostitution, Forgery/BadChecks,* and *Rape* were > 1.0 logit, i.e., easier for adults to endorse. This finding mirrors the gender DIF findings because adult females tended to endorse *Prostitution, Forgery/BadChecks,* and *Rape*.

Therefore, adults and females were driving the DIF on these items. Please see Riley, Conrad, Chan, and Dennis (2007) for a complete treatment of this issue.

Race DIF. In Figure 3, we displayed DIF among the four racial/ethnic groups. Of note was that *GangMember* and *ThreatenSomeoneWithGunOrKnife* tended to be more difficult for Caucasians to endorse. African Americans tended to endorse *Forced Sex* more than the other groups.

Primary Drug Severity DIF. With six different types of drugs, it appears that opiates and cocaine have the most severe effects. Specifically, users of cocaine and opiates found it significantly easier to endorse *Prostitution* and *Forgery/BadChecks*.

Figure 2. CVS Gender DIF

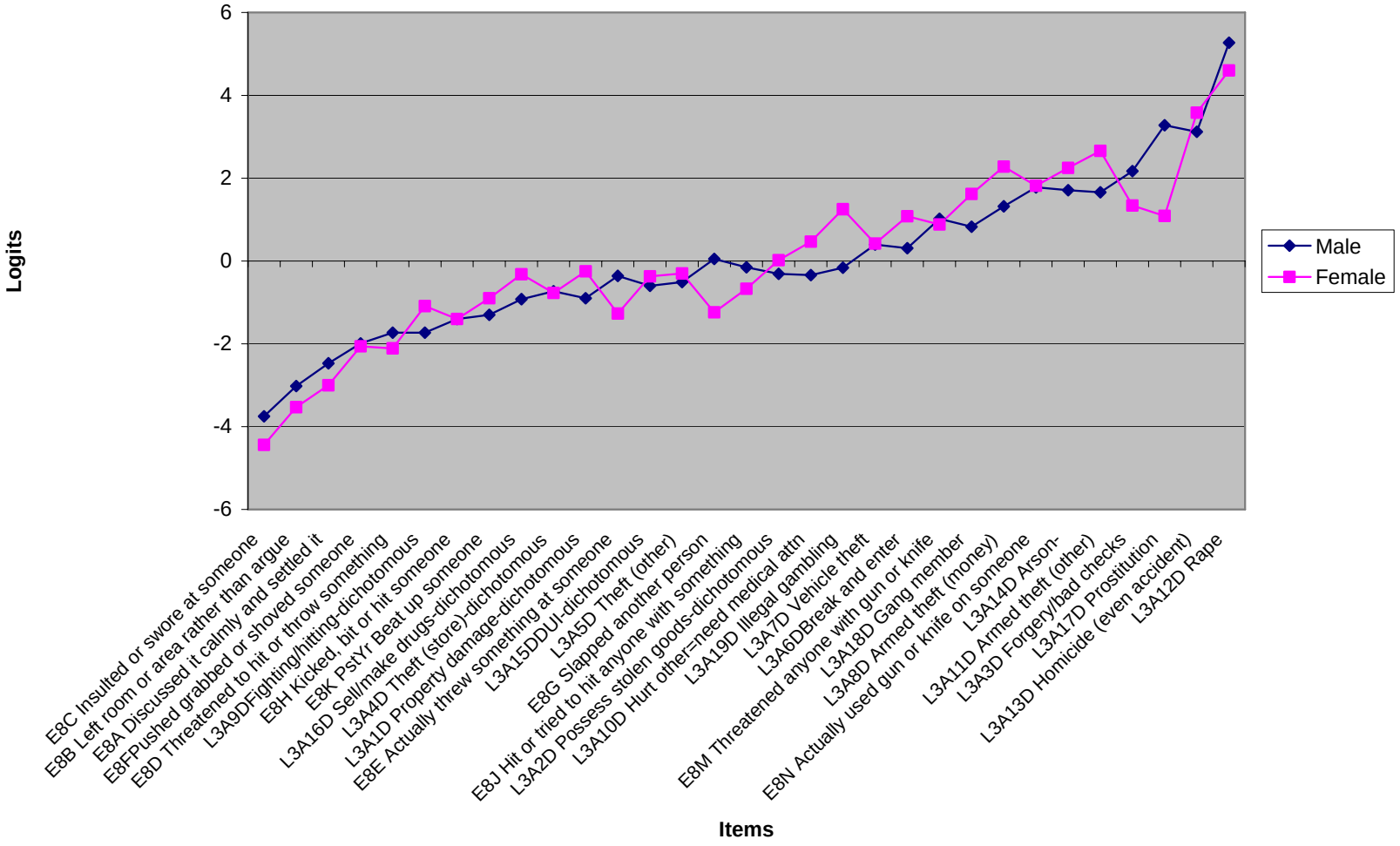


Figure 3. CVS Age DIF

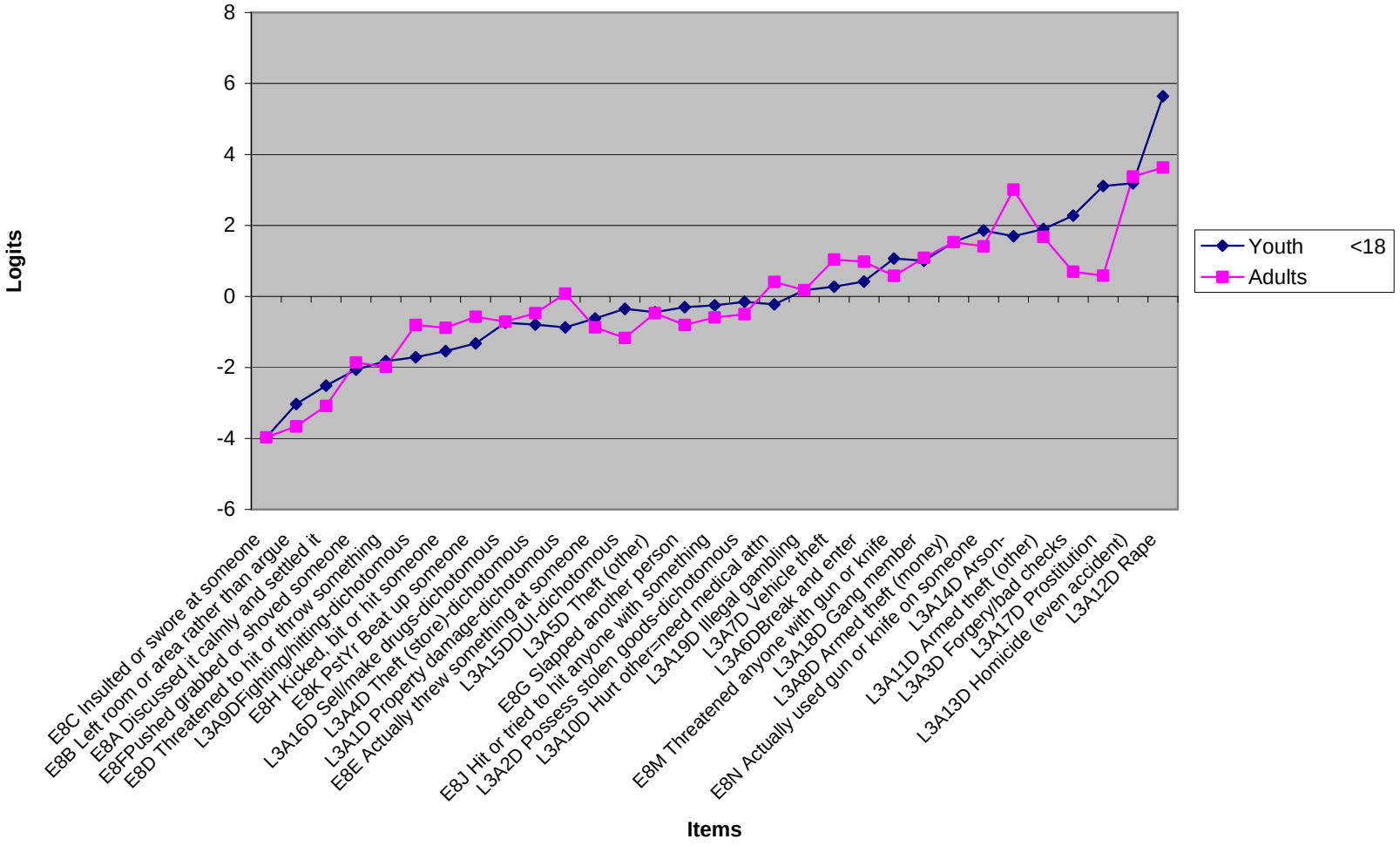


Figure 4. CVS Race DIF

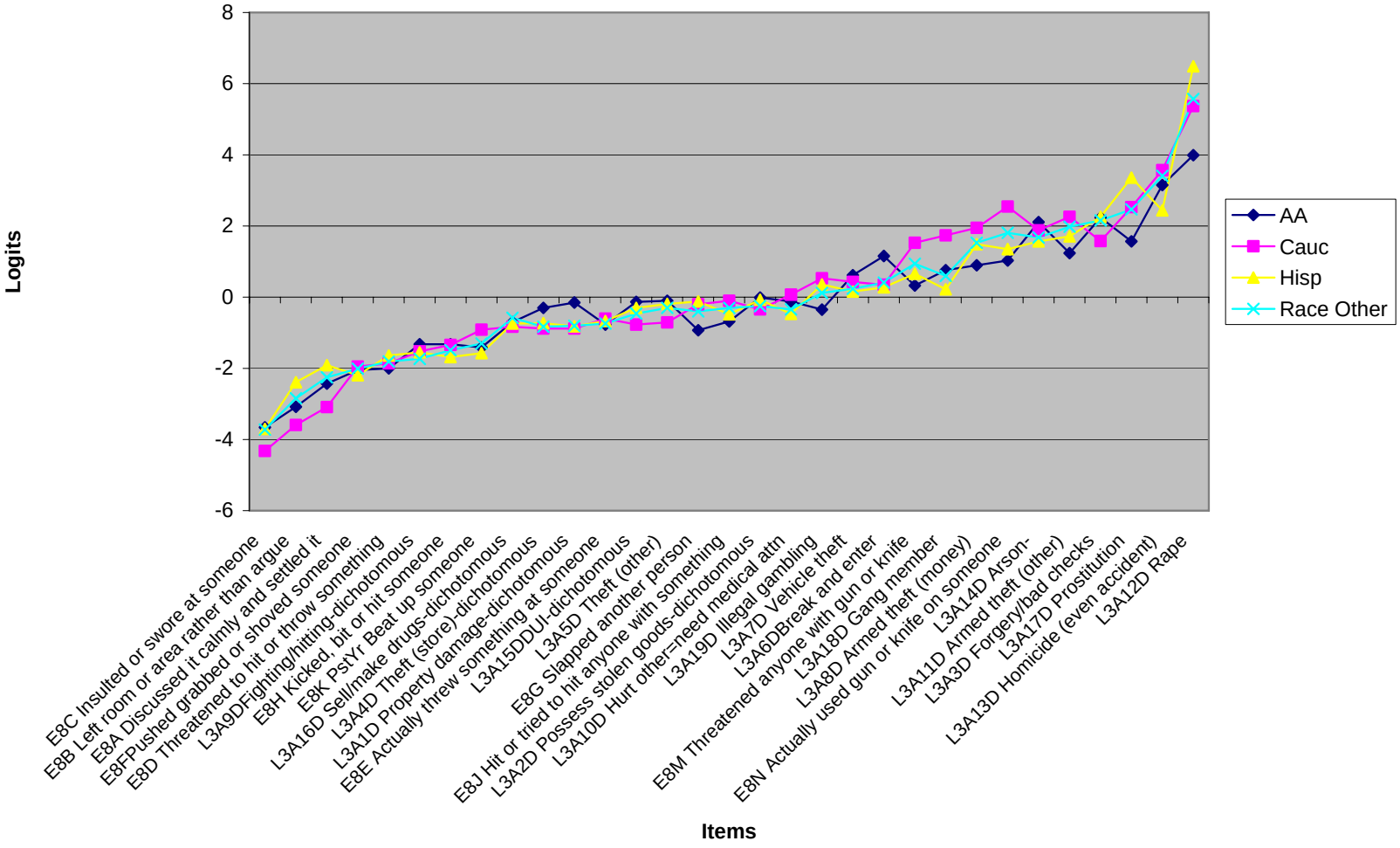


Figure 5. CVS Drug Severity DIF

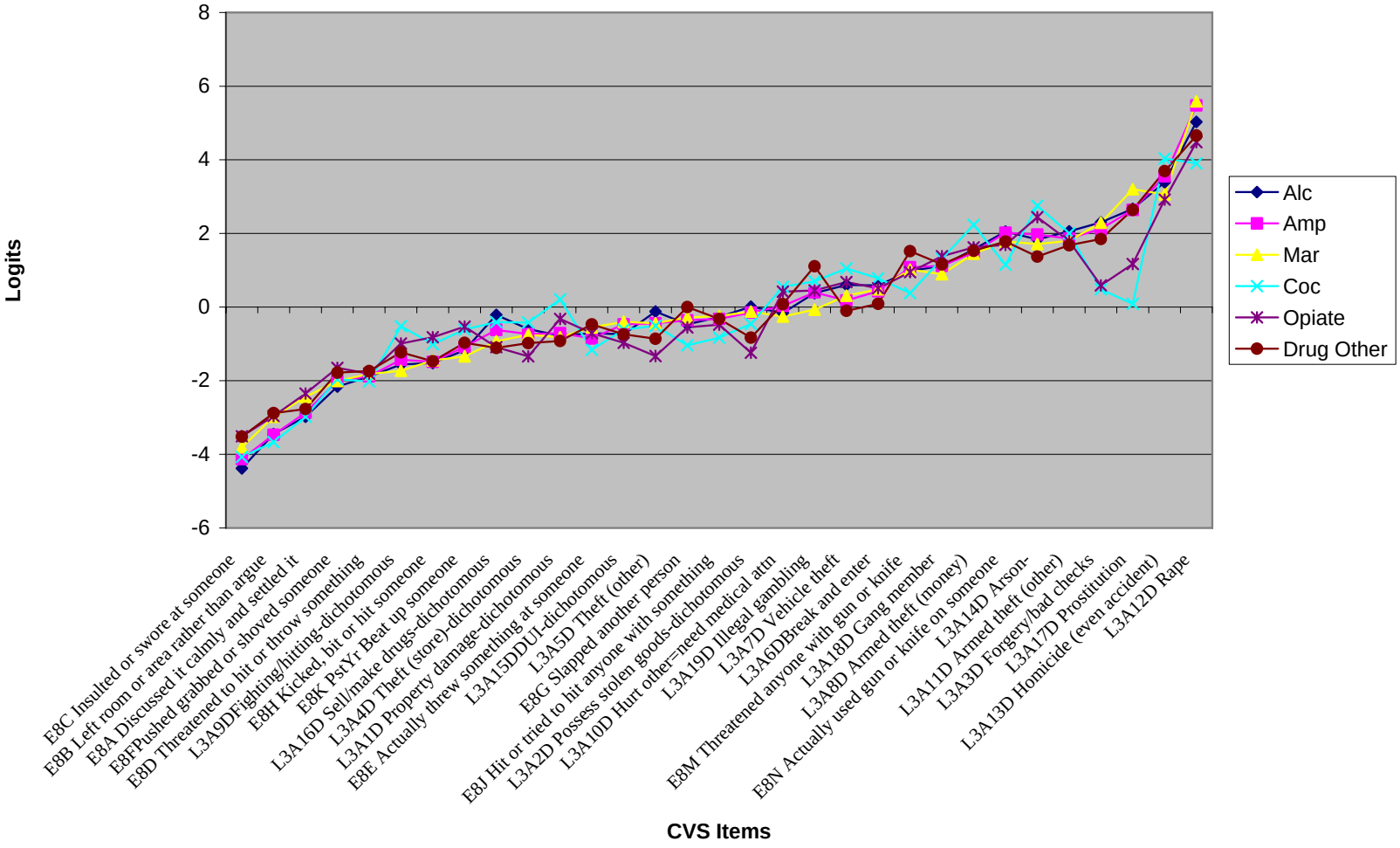


Table 1. CVS Item Measures by Demographic Group (Items listed in Severity Order)

CVS Item Labels	Male	Female	Youth <18	Adults	AA	Cauc	Hisp	Race Other	Alc	Amp	Mar	Coc	Opiate	Drug Other
E8C Insulted or swore at someone	-3.75	-4.44	-3.97	-3.97	-3.66	-4.32	-3.7	-3.72	-4.38	-4.13	-3.81	-4.07	-3.51	-3.52
E8B Left room or area rather than argue	-3.02	-3.53	-3.03	-3.66	-3.08	-3.59	-2.39	-2.84	-3.45	-3.47	-2.98	-3.67	-2.96	-2.88
E8A Discussed it calmly and settled it	-2.47	-3	-2.51	-3.08	-2.43	-3.09	-1.91	-2.25	-2.98	-2.87	-2.44	-2.98	-2.35	-2.77
E8FPushed grabbed or shoved someone	-1.99	-2.06	-2.06	-1.86	-2.05	-1.95	-2.19	-2.01	-2.16	-2.01	-2.01	-1.96	-1.65	-1.78
E8D Threatened to hit or throw something	-1.73	-2.11	-1.82	-1.98	-2	-1.86	-1.64	-1.81	-1.89	-1.88	-1.82	-2.02	-1.81	-1.74
L3A9DFighting/hitting-dichotomous	-1.73	-1.09	-1.71	-0.8	-1.32	-1.52	-1.55	-1.73	-1.56	-1.43	-1.73	-0.52	-0.99	-1.23
E8H Kicked, bit or hit someone	-1.4	-1.4	-1.54	-0.88	-1.32	-1.34	-1.68	-1.49	-1.52	-1.49	-1.45	-1.01	-0.82	-1.47
E8K PstYr Beat up someone	-1.3	-0.9	-1.32	-0.57	-1.41	-0.91	-1.57	-1.3	-1.17	-1.08	-1.34	-0.61	-0.53	-0.97
L3A16D Sell/make drugs-dichotomous	-0.92	-0.32	-0.74	-0.71	-0.7	-0.83	-0.74	-0.58	-0.21	-0.62	-0.93	-0.41	-1.09	-1.11
L3A4D Theft (store)-dichotomous	-0.73	-0.77	-0.79	-0.47	-0.3	-0.88	-0.73	-0.83	-0.6	-0.73	-0.75	-0.42	-1.34	-0.98
L3A1D Property damage-dichotomous	-0.9	-0.25	-0.87	0.08	-0.15	-0.88	-0.82	-0.8	-0.78	-0.71	-0.82	0.21	-0.32	-0.92
E8E Actually threw something at someone	-0.36	-1.27	-0.62	-0.87	-0.77	-0.6	-0.67	-0.74	-0.71	-0.85	-0.55	-1.16	-0.71	-0.47
L3A15DDUI-dichotomous	-0.6	-0.37	-0.35	-1.17	-0.13	-0.77	-0.3	-0.46	-0.74	-0.47	-0.4	-0.61	-0.97	-0.74
L3A5D Theft (other)	-0.51	-0.3	-0.44	-0.47	-0.1	-0.71	-0.19	-0.3	-0.12	-0.44	-0.44	-0.51	-1.33	-0.86
E8G Slapped another person	0.05	-1.24	-0.3	-0.8	-0.93	-0.2	-0.12	-0.4	-0.47	-0.36	-0.28	-1.04	-0.55	0
E8J Hit or tried to hit anyone with something	-0.15	-0.67	-0.25	-0.59	-0.68	-0.1	-0.48	-0.3	-0.27	-0.32	-0.24	-0.83	-0.48	-0.32
L3A2D Possess stolen goods-dichotomous	-0.31	0.02	-0.15	-0.5	-0.01	-0.34	-0.07	-0.25	0.01	-0.15	-0.12	-0.46	-1.24	-0.83
L3A10D Hurt other=need medical attn	-0.34	0.47	-0.22	0.41	-0.12	0.07	-0.47	-0.34	-0.16	0.03	-0.26	0.55	0.42	0.08
L3A19D Illegal gambling	-0.16	1.25	0.18	0.18	-0.35	0.53	0.36	0.12	0.38	0.4	-0.07	0.7	0.45	1.11
L3A7D Vehicle theft	0.4	0.42	0.28	1.04	0.62	0.43	0.16	0.24	0.6	0.18	0.31	1.05	0.68	-0.1
L3A6DBreak and enter	0.31	1.08	0.42	0.98	1.16	0.35	0.28	0.41	0.59	0.43	0.48	0.78	0.51	0.09
E8M Threatened anyone with gun or knife	1.02	0.88	1.07	0.58	0.33	1.53	0.66	0.94	0.98	1.08	1.02	0.38	0.95	1.52
L3A18D Gang member	0.83	1.62	1.01	1.09	0.76	1.74	0.22	0.6	1.12	1.12	0.89	1.31	1.38	1.16
L3A8D Armed theft (money)	1.32	2.28	1.53	1.53	0.9	1.95	1.49	1.53	1.57	1.48	1.45	2.23	1.62	1.53
E8N Actually used gun or knife on someone	1.78	1.82	1.86	1.41	1.03	2.55	1.35	1.81	2.05	2.02	1.78	1.15	1.69	1.78
L3A14D Arson-	1.71	2.25	1.7	3.01	2.11	1.88	1.57	1.68	1.84	1.97	1.71	2.75	2.44	1.37
L3A11D Armed theft (other)	1.66	2.66	1.9	1.68	1.24	2.26	1.71	1.98	2.06	1.86	1.8	2	1.79	1.68
L3A3D Forgery/bad checks	2.17	1.34	2.28	0.7	2.24	1.58	2.26	2.16	2.3	2.12	2.29	0.48	0.59	1.85
L3A17D Prostitution	3.28	1.09	3.11	0.59	1.57	2.53	3.36	2.47	2.67	2.64	3.2	0.09	1.17	2.64
L3A13D Homicide (even accident)	3.12	3.58	3.19	3.38	3.15	3.57	2.44	3.41	3.39	3.55	3.06	4.04	2.92	3.69
L3A12D Rape	5.27	4.6	5.64	3.63	3.99	5.37	6.49	5.57	5.03	5.48	5.59	3.9	4.48	4.66

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 serious crimes. 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 < or = 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). In the figures, the solid lines represent the actual item response means for each fit group (i.e., fit group item means) and the dashed lines represent the item means over all persons (i.e., overall item means). The red vertical lines indicate the difference between overall item means and fit group item means. Recall that the items are dichotomous with 0/1 categories).

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.

In Figure 8, the L/M infit and HI outfit group (7%) is called Atypical Type 1, where the person measure may underestimate severity slightly since these tend to be people who report serious crimes (higher criminality) but are unexpectedly low or on target on only two items. These are the two most misfitting items, i.e., *E8A Discussed it calmly and settled it* and *E8B Left room or area rather than argue*. If these two items were deleted, most of the people in this group would probably not be misfitting. This is further support for dropping these two items from the person measures.

In Figure 9, the HI infit and L/M outfit group (2%) is called Atypical Type 2 and tends to be persons with moderate measures whose actual item means are unexpectedly lower on the less serious Conflict Tactic Scale items and unexpectedly higher on property crimes, fighting/hitting, and drug crimes. They are only average on physical violence and interpersonal crime. These persons should be red-flagged since they have this unusual response pattern.

In Figure 10, the HI infit and HI outfit group (12%) called Atypical Type 3 tend to be persons who are high in criminality but have an overall measure that may underestimate severity. They tend to be persons who endorse the more serious property, interpersonal, and drug crimes and not endorse the less serious Conflict Tactic Scale items. These persons misfit because they tend to endorse most of the higher seriousness crimes but do not tend to endorse some of the lower risk items, such as discussing things calmly, leaving the room, or having oral arguments to resolve conflicts. Their measures should also be red-flagged and viewed with caution since they will tend to be lower than they should be. They may have a deficit in non-violent conflict resolution skills.

In summary, the pattern that is most deceptive is Atypical Type 3 since the person measures will tend to be substantially lower than they should be based on the persons' lack of endorsement of Conflict Tactic Scale items and their endorsement of serious crimes. Atypical Type 2 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 more serious crimes. Atypical 1 has only minor deviations from expectations and therefore this pattern has little clinical significance.

Figure 6. CVS Fit Groups

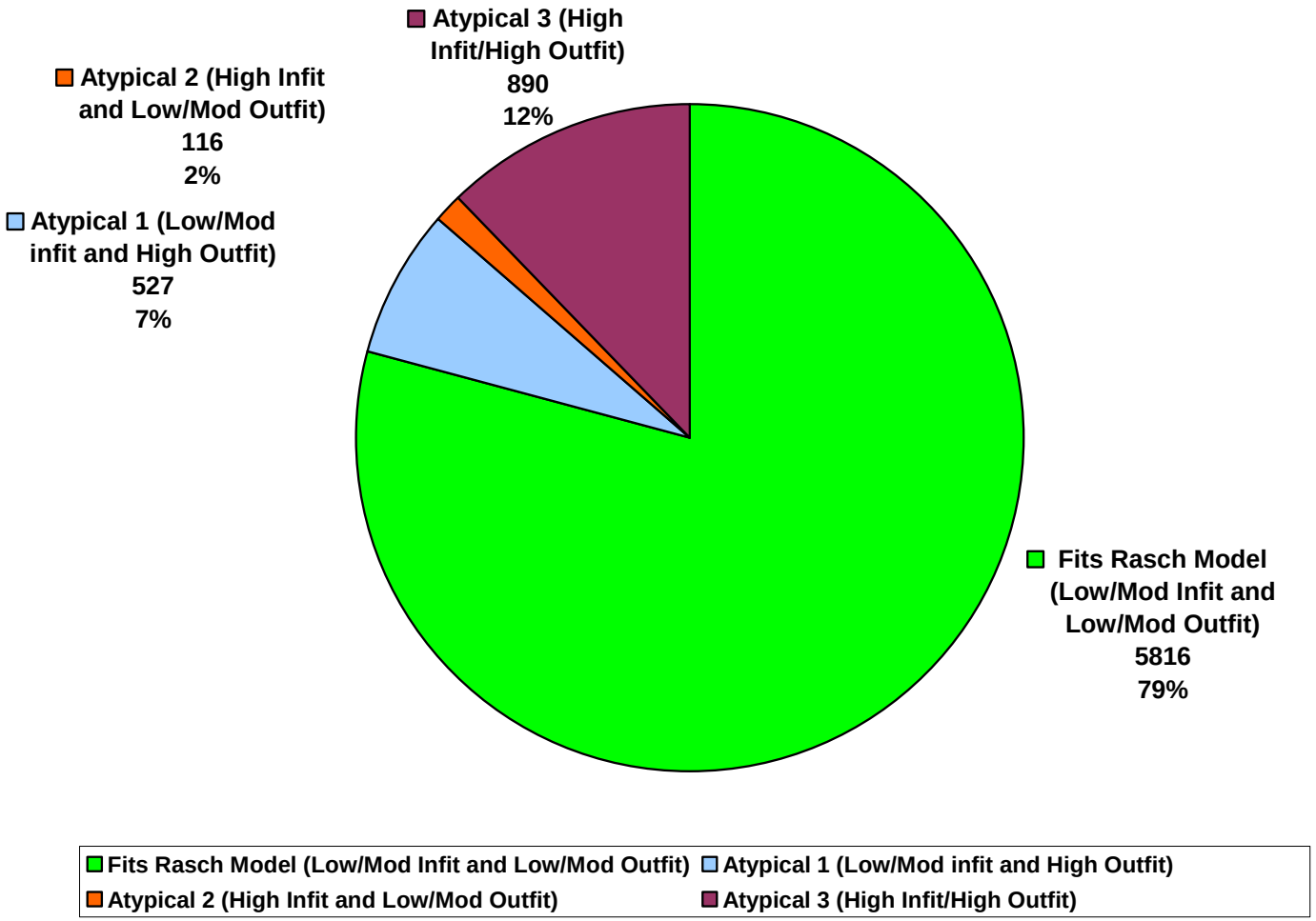


Figure 7. CVS Overall Item Mean vs. Fit Group Item Mean
Typical Group: Fits the Rasch Model (LMLM) (n=5816; 79%)
(Low/Mod on Infit and Low/Mod on Outfit)

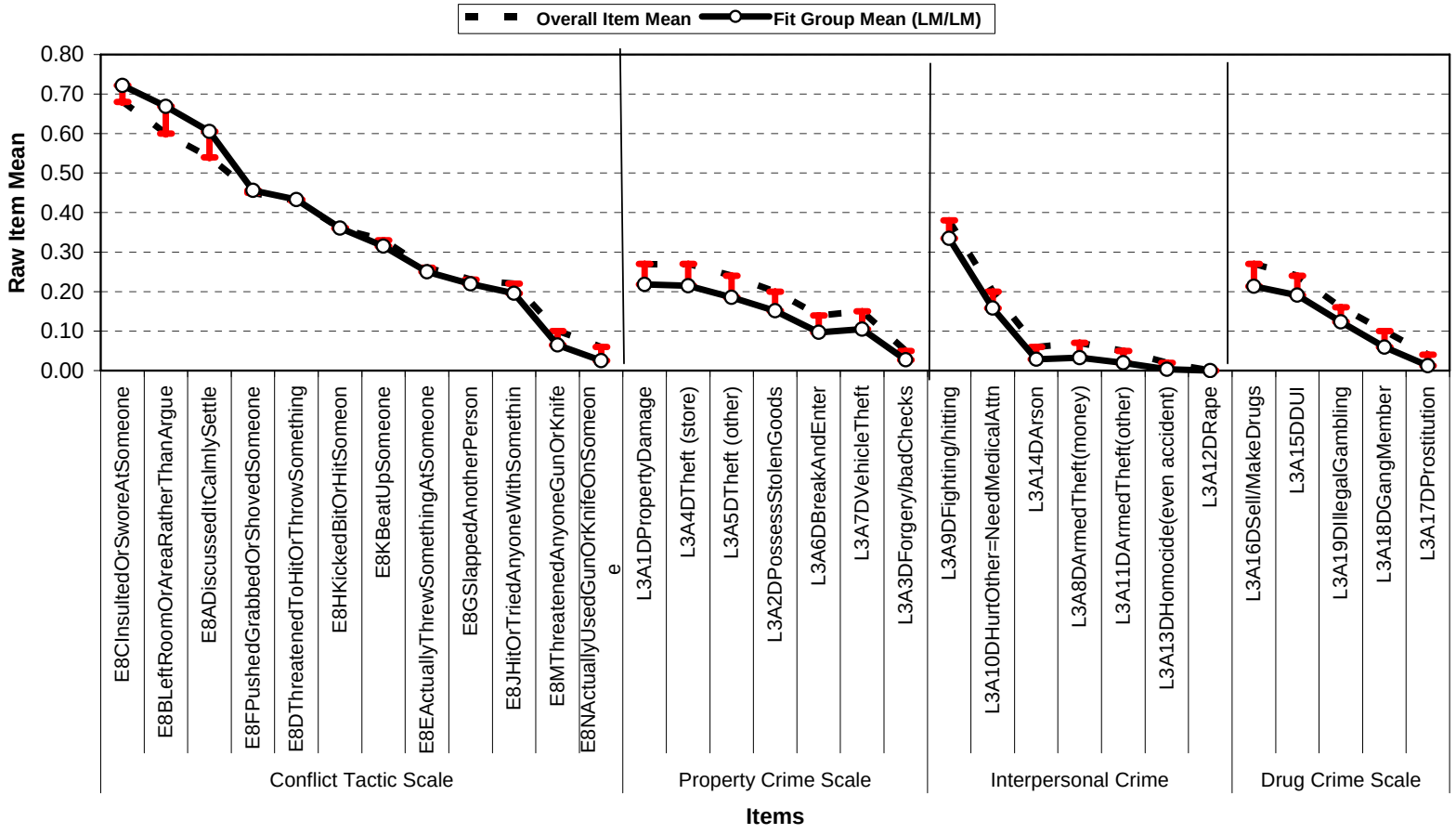


Figure 8. CVS Overall Item Mean vs. Fit Group Item Mean
Aytpical Type 1 with Higher Criminality, but Lower on Two Low Seriousness Items
(n=527; 7%, Low/Mod Infit and High Outfit)

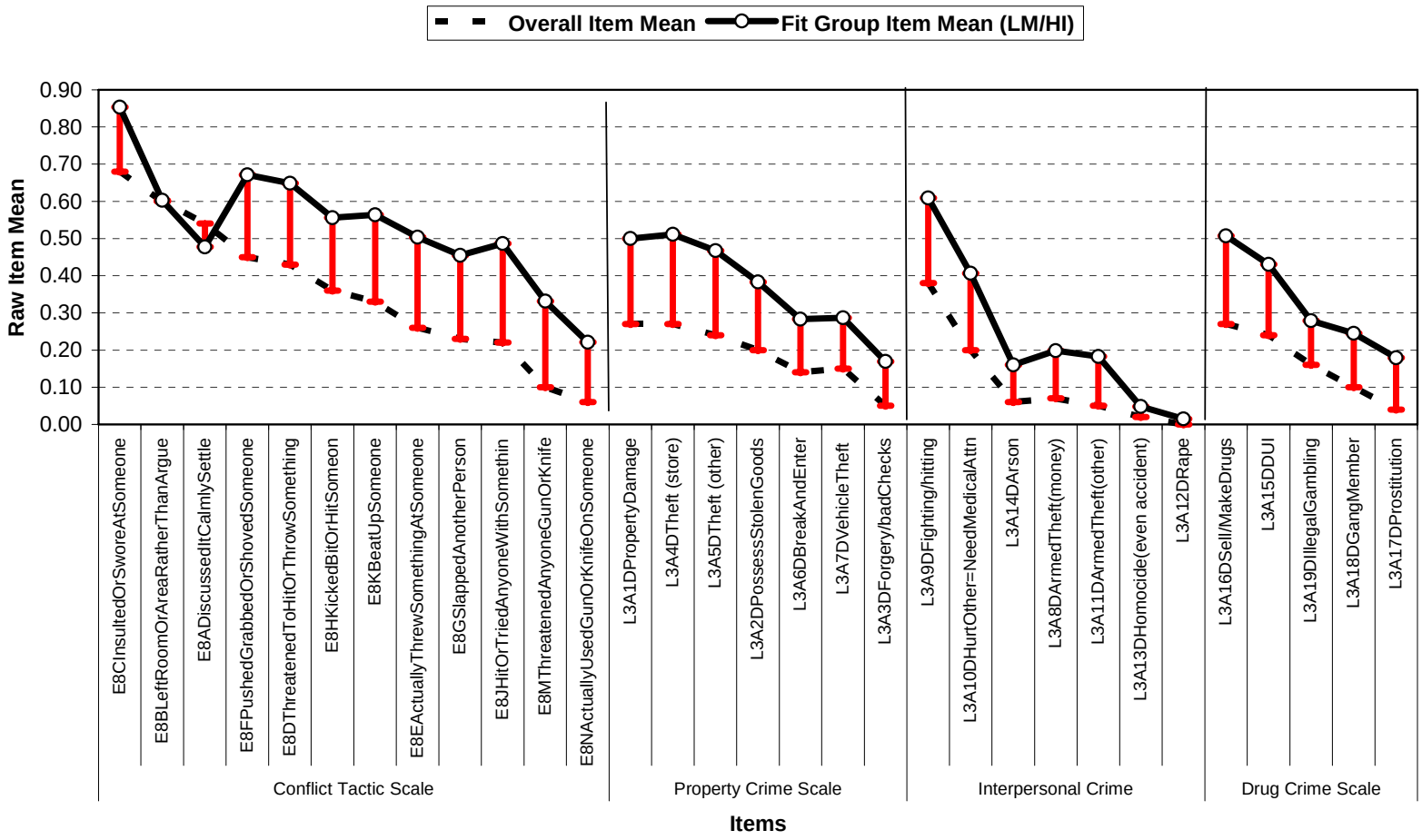


Figure 9. CVS Overall Item Mean vs. Fit Group Item Mean:
Aytpical Type 2 with Lower Conflict Tactics and Higher Drug and Property Crime and Fighting
(n=116; 2%)
High Infit and Low/Mod Outfit

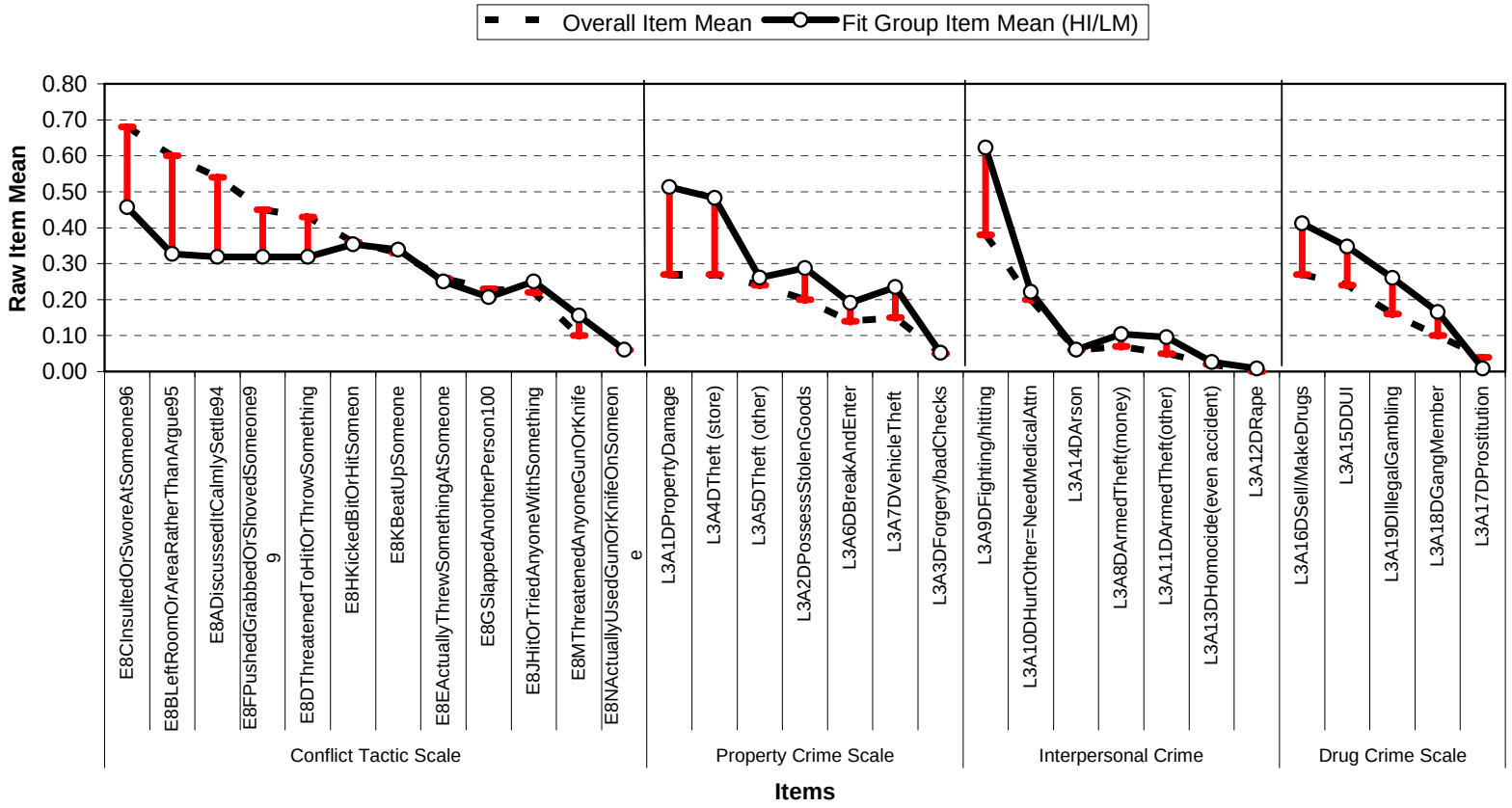
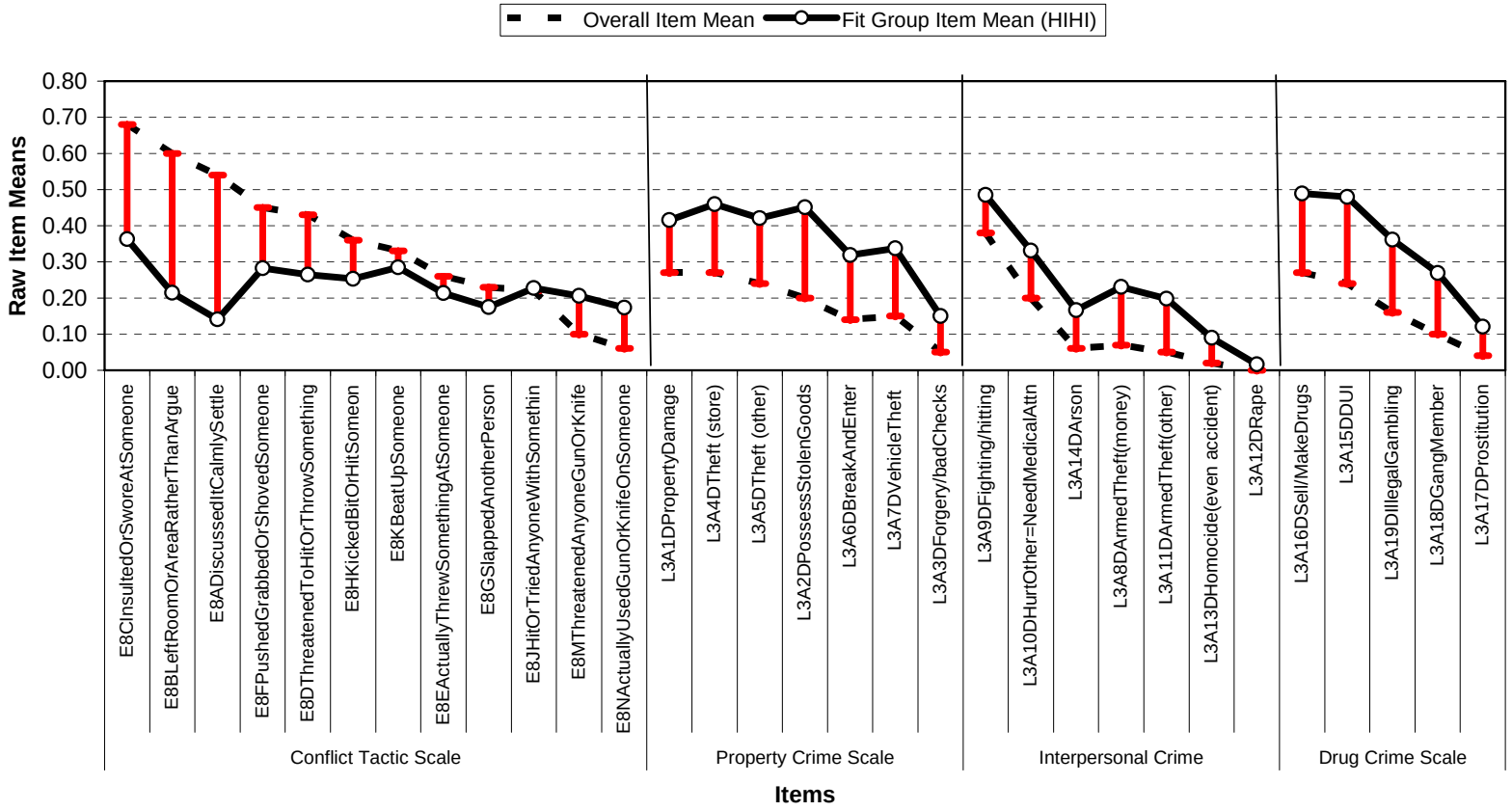


Figure 10. CVS Expected Item Mean vs. Actual Item Mean
Aytpical Type 3 with Lower on Conflict Tactics and Higher on Criminality, More Extreme than Type 2 (n=890; 12%)
High Infit and High Outfit



Summary

Both items and scales form the theoretically expected hierarchies with a person internal consistency reliability of .82 and an item reliability of 1.00. Of the 31 items in the *CVS*, significant DIF (i.e., > .5 SD = .94 logits) occurred in 4 items for males vs. females, 5 items for youth vs. adults, 2 items for race when using Caucasians as the reference group, and 2 items for primary substances when using alcohol as the referent. The persons' responses generally conformed to the expectations of the Rasch model with about 79% of the sample providing expected responses. Two items misfit the CVS both psychometrically and logically i.e., *Discussed It Calmly and Settled the Disagreement* and *Left the Room or Area Rather Than Argue*. They seem to have little to do with the construct of crime and violence.

Recommended Actions and Deliberations

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

- Two items misfit the CVS both psychometrically and logically. These items, *Discussed It Calmly and Settled the Disagreement* and *Left the Room or Area Rather Than Argue,* seem to have little to do with the construct of crime and violence.
- Consider separate CVS measures for men and women and for youth and adults based on the DIF analysis. The items *Prostitution* and *Forgery/BadChecks* are much easier for females, adults, and cocaine/opiate users to endorse.
- Because the sample is predominantly young males who do not tend to endorse *Prostitution* and *Forgery/BadChecks*, these items appear to be much more severe/more rare crimes than they should be. In other words, if this were a sample of women, these items would be much lower in the seriousness hierarchy. For a complete treatment of this latter issue, please contact the authors for a manuscript by Conrad, Chan, Riley, Conrad, and Dennis (2007).
- Two groups, Atypical Types 2 and 3, tend to have measures that underestimate the seriousness of their criminality. We recommend flagging these two groups for clinicians in the evaluation/validity concerns section of the GAIN Recommendation and Referral Summary (GRRS) as:
    - Atypical Type 2 response pattern on the Crime and Violence Scale (Relative to the total score, lower than expected on the less serious conflict tactic symptoms and higher than expected on property crimes, fighting/hitting, and drug crimes)
    - Atypical Type 3 response pattern on the Crime and Violence Scale (Relative to the total score, lower than expected on the less serious conflict tactic symptoms and higher than expected in criminality)
- More work on construct validity would be helpful to understand these fit groups better and to ensure proper interpretation of the measures.

References

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.

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. (2008). *Winsteps Rasch measurement computer program* (Version 3.64.2). Available from http://www.winsteps.com/

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 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.

Riley, B.B., Conrad, K.J., Chan, Y.F. Conrad, K.M., & Dennis, M.L. Gender differences in items measuring crime and violence: A Rasch analysis. Working Paper, Chestnut Health Systems.

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.

March 20, 2009