Online Supplement with Addendum

Conjectures and Refutations in Cognitive Ability Structural Validity Research: Insights from Bayesian Structural Equation Modeling

By S. C. Dombrowski et al. (2025)

Journal of School Psychology

Table A1

Descriptive Statistics for WISC-V Subtest and Index Scores

Score	*N*	Mean	SD	Skewness	Kurtosis
Block Design	710	8.6	3.5	+0.23	-0.02
Similarities	710	9.2	3.4	-0.01	-0.17
Matrix Reasoning	710	9.1	3.4	-0.07	-0.30
Digit Span	710	8.0	3.1	+0.22	+0.26
Coding	710	7.3	3.3	+0.06	-0.36
Vocabulary	710	8.9	3.6	+0.06	-0.58
Figure Weights	710	9.6	3.2	+0.05	-0.29
Visual Puzzles	710	9.7	3.4	+0.02	-0.34
Picture Span	710	8.6	3.3	+0.10	-0.30
Symbol Search	710	8.1	3.3	+0.00	-0.01
VCI	702	94.8	18.0	-0.03	-0.25
VSI	703	95.3	17.9	0.17	-0.03
FRI	702	96.2	17.3	-0.03	-0.44
WMI	703	90.1	16.2	0.14	-0.16
PSI	703	87.0	17.1	-0.10	-0.02
FSIQ	670	90.8	17.7	+0.04	-0.27

Note. VCI = Verbal Comprehension Index; VSI = Visual Spatial Index; FRI = Fluid Reasoning Index; WMI = Working Memory Index; PSI = Processing Speed Index; FSIQ = Full Scale Intelligence Quotient.

Table A2

Covariance and Correlations Matrices

Covariance Matrix

	BD	SI	MR	DS	CD	VO	FW	VP	PS	SS
BD	11.922									
SI	6.28	11.724								
MR	7.2	5.929	11.634							
DS	5.231	6.023	5.536	9.596						
CD	4.663	3.974	4.421	4.425	10.784					
VO	7.035	9.336	6.442	6.408	4.702	12.857				
FW	7.177	6.137	6.785	5.001	3.782	6.673	10.17			
VP	8.647	6.573	7.254	5.427	4.625	7.188	7.062	11.381		
PS	5.007	4.916	5.115	5.783	4.12	6.025	4.683	5.279	10.689	
SS	5.323	4.468	4.734	4.502	6.704	5.094	3.903	5.272	4.214	10.841

Correlation Matrix

	BD	SI	MR	DS	CD	VO	FW	VP	PS	SS
BD	1									
SI	0.531	1								
MR	0.611	0.508	1							
DS	0.489	0.568	0.524	1						
CD	0.411	0.353	0.395	0.435	1					
VO	0.568	0.76	0.527	0.577	0.399	1				
FW	0.652	0.562	0.624	0.506	0.361	0.584	1			
VP	0.742	0.569	0.63	0.519	0.417	0.594	0.656	1		
PS	0.444	0.439	0.459	0.571	0.384	0.514	0.449	0.479	1	
SS	0.468	0.396	0.422	0.441	0.62	0.431	0.372	0.475	0.391	1

Table A3

Mplus Input Code

```
Title:     BSEM WISC-5 Clinical Sample 10 Subtest
        !Model 5 Higher Order OR 5 Bifactor Model (VCI, VZI FRI, WMI, PSI)
data:
file is "WISCV.txt";
VARIABLE:
NAMES ARE bd si mr ds cd vo fw vp ps ss;
USEVARIABLES ARE bd si mr ds cd vo fw vp ps ss ;
define:
        standardize bd si mr ds cd vo fw vp ps ss  ;
analysis:
estimator = bayes;
!estimator = mlm;
proc = 2;
fbiter = 100000;
chain = 2
stvalues = ml;
Kolmogorov = 100;

MODEL:
! 5  Higher Order
vci by si vo ;
vzi by bd vp;
fri by mr fw;
wmi by ds ps ;
psi by cd ss;
g by vci vzi fri wmi psi ;

! For bifactor model, cross loads and correlated residuals code remains the same but replace with the followig code:
! 5 Bifactor

!vci by si* vo (1);
!vci@1;

!vzi by bd* vp (2);
!vzi@1;

!fri by mr* fw (3);
!fri@1;

!wmi by ds* ps (4);
!wmi@1;

!psi by cd* ss (5);
!psi@1;

!g by si* vo bd vp mr fw ds ps cd ss ;
!g@1;

!vci with vzi-g@0;
!vzi with fri-g@0;
!fri with wmi-g@0;
!wmi with psi-g@0;
!psi with g@0;


!cross-loadings:
!   Saying *0 gives a zero mean start value for the parameter.
!   The prior with mean zero and small variance (e.g., .01) is then applied during the computations.

vci by bd*0 vp*0 mr*0 fw*0 ds*0 ps*0 cd*0 ss*0(xload1-xload8);

vzi by si*0 vo*0 mr*0 fw*0 ds*0 ps*0 cd*0 ss*0 (xload9-xload16);

fri by si*0 vo*0 bd*0 vp*0 ds*0 ps*0 cd*0 ss*0 (xload17-xload24);

wmi by si*0 vo*0 bd*0 vp*0 mr*0 fw*0 cd*0 ss*0 (xload25-xload32);

psi by si*0 vo*0 bd*0 vp*0 mr*0 fw*0 ds*0 ps*0 (xload33-xload40);


!correlated error terms
```

```
      bd-ss(p1-p10);

!bd-ss are the residual variances of the subtests.
!The IW prior is assigned for the residual variance using b1-b10 as described below.

!These are the subtest covariances lower-triangular elements taken row-wise

      bd-ss with bd-ss (c1-c45);
      vci-psi(a1-a5);
      vci-psi with vci-psi(b1-b10);

!b1=error covariance; a1=error variance; IW defined below and assigned as noted above
!b1 a1
!b2 b1 a2
!b3 b1 b2 a3
!b4 b1 b2 b3 a4

      model priors:

!Create the residual variance-covariance matrix (by default, residuals are independent, thus not correlated).
!This will allow to estimate them because otherwise they wouldn't be part of the model.
!There will be 45 covariances and 10 residual variances in total. Forumla: n(n-1)/2 for # of covar.
!Name the variances from b1-b10 and the covariances from c1-c45


!This is the approach described in Muthen & Asparouhov (2012).

!16 = 10 PARAM + 6 !Assign an inverse-wishart prior to the residual variances
!(following Muthen and Asparouhov instructionsn see page 14-15)
!df=# of parameters +6; where IW (I, df)

!p1-p10~IW(1,16);

!##################################################################################################################
!Steps to conduct corr residuals based on Asparouhov et al (2015). See Asparouhov et al (2015) for technical description in appendix

!Steps to conduct correlated residuals.
!1) Run xloads only analysis.
!2) Obtain residual var from each indicator.
!3) The forumla is IW(residual var*df, df) or (.282*100, 100).!Df is obtained from Asparouhov et al. (2015) where
!df is chosen based upon sample size such that N=10,000 then df=1,000. N=500 the df =100.
!It takes an additional step and specification as noted above and reference to their response to Stromeyer et al (2015) to obatin df.
!In this example since N=710, this is about ~500 so d =100 for this formula.

            p1~IW(28.2, 100);
            p2~IW(27.3, 100);
            p3~IW(43.5, 100);
            p4~IW(31.2, 100);
            p5~IW(46.9, 100);
            p6~IW(20.7, 100);
            p7~IW(37.9, 100);
            p8~IW(26.2, 100);
            p9~IW(51.8, 100);
            p10~IW(26.9, 100);

!Assigns IW prior to latent group residual variances.
            a1-a5~IW(1,11);

!Assign another kind of inverse-wishart prior to the subtest covariances (Asparouhove &Muthen, 2015)
            c1-c45~IW(0,100);

!Assigns an IW to latent group factor covariances based on Muthen & Asparouhov (2012)
            b1-b10~IW(0, 11);

!Assigns prior variance to the xloads

            xload1-xload40~N(0, .01); !Assigns prior variance of .01 to the xloads


            output:
            tech1 tech8 stdy svalues;

            plot:
            type = plot2;
```

Table A4

Five Factor Higher Order Model Loading Estimates

	No priors	Xloads Only	Xloads & Corr Resid (Subtests)	Xloads & Corr Resid (Subtests & Group Factors)
Verbal				
SI	.85	.81	.65	.77
VO	.89	.78	.89	.86
Visual Spatial				
BD	.46	.94	.89	.85
VP	.41	.89	.94	.82
Fluid Reasoning				
MR	.78	.77	.80	.75
FW	.80	.84	.74	.75
Working Memory				
DS	.80	.74	.74	.79
PS	.71	.53	.53	.66
Processing Speed				
CD	.75	.72	.82	.66
SS	.82	.80	.71	.88
General				
VCI	.84	.78	.77	.85
VSI	.92	.95	.96	.91
FRI	.97	.98	.99	.90
WMI	.86	.75	.78	.88
PSI	.69	.66	.66	.76

Note. Xloads = Cross-loadings, Corr resid = Correlated residual. Correlated residual between FRI and VSI (.432) significant at $p < .05$

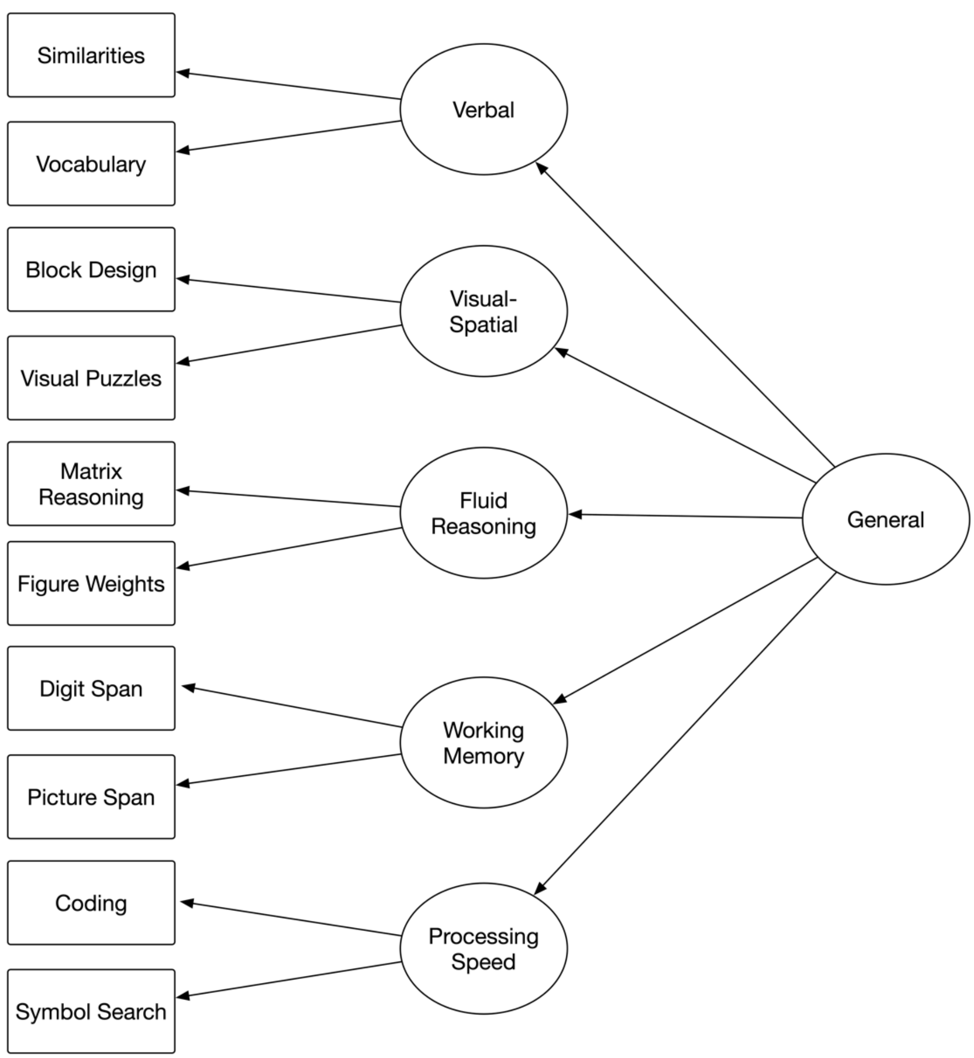
Addendum

Higher-Order vs. Bifactor Models: Distinctions with a *Slight* Difference

When modelling the factor structure of IQ tests, researchers generally consider two different structural models. The first is the higher-order (second-order) model where the general (second-order) factor is fully mediated by first-order group factors in influencing subtests indicators (measured variables) below the group factors. This model is depicted in Figure 1 below using the WISC-V ten subtest primary battery.

Figure 1

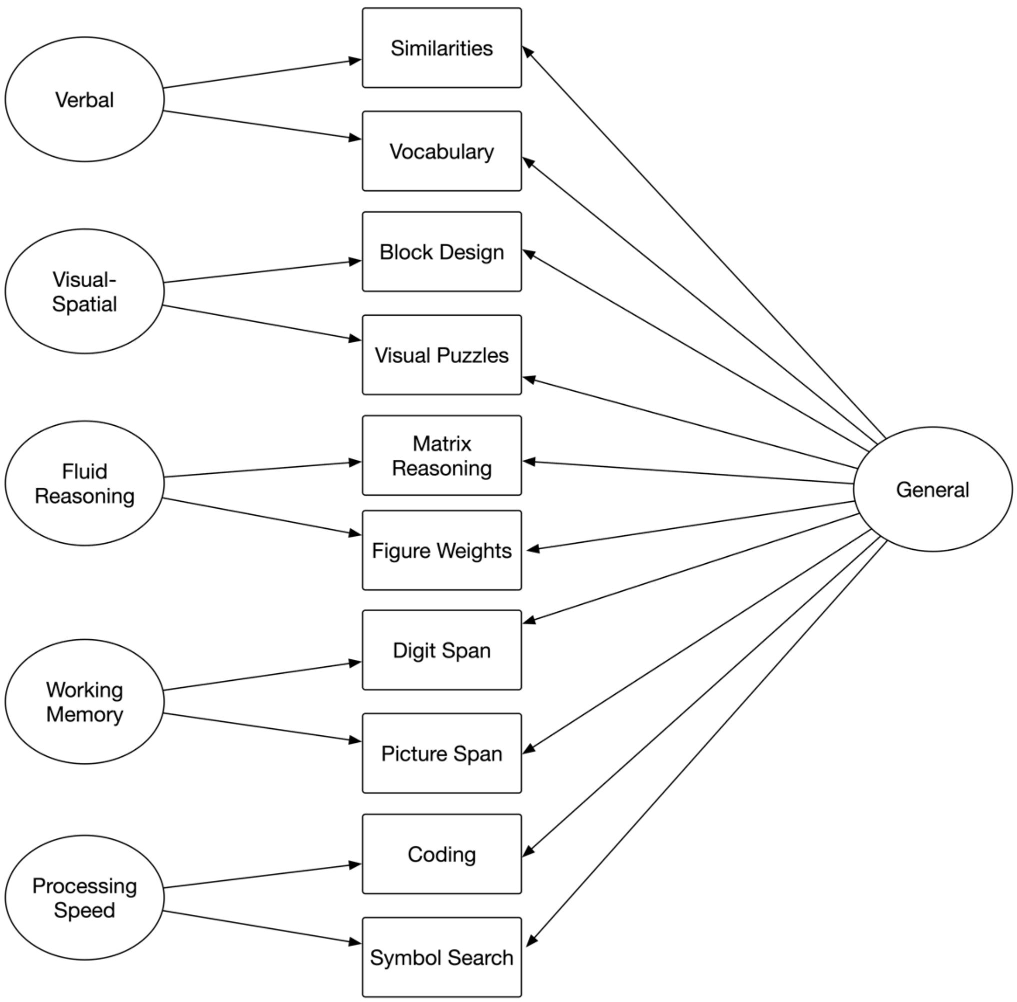
Higher Order Model for the WISC-V



The second is the bifactor model where the general factor and group factors simultaneously have direct influences on individual subtest indicators (measured variables). This is depicted in Figure 2 below using the WISC-V ten subtest battery.

Figure 2

Bifactor Model for the WISC-V



While both models acknowledge the existence of the general factor (*g*) on intelligence tests, the higher-order model conceptualizes the general factor as superordinate to the group factors and subtests with first-order group factors between the general factor and the individual subtests (measured variables). In other words, the general factor has no direct influence on the individual subtests as its influence is fully mediated by the first-order group factors. Conversely, the bifactor model conceptualizes the general factor as a breadth factor and assumes that *g* and the group factors have simultaneous direct effects on the subtests. For a more technical discussion of this topic please see Beaujean (2015), Canivez (2016), and Gignac (2008). From a practical, interpretive perspective, the bifactor model provides partitioned variance to determine the relative influence of the general factor versus group factors which has implications for their use in interpretive methods which stress primary interpretation of group factor indices (e.g., PSW). This may be accomplished with a higher order model (as per Keith & Reynolds, 2018) but the merits of variance partitioning using the HO model[1] have not been thoroughly investigated so this practice should be considered experimental until further verification ensues.

From a theoretical perspective the higher order model's conceptualization of *g* is akin to looking at the shadow of a person standing next to a street light versus looking directly at the person (bifactor) to ascertain how tall they are. The higher-order model was explicated by Thurstone (1947) and is considered a bottom-up (American model) where group factors are prioritized (see Beaujean & Benson, 2019 for a discussion). Spearman (1904) was among the first to discuss *g* where he posited a two-factor theory for the construct. With Spearman's model (i.e., the British approach) interpretation of the general factor is prioritized and group/specific factors are regarded as largely a statistical nuisance. However, in some of his later writings, Spearman regarded group/specific factors with greater importance. In fact, one of Spearman's post-doctoral fellows, Karl Holzinger, elaborated on two-factor theory culminating in the development of what later became known as the bifactor model (Holzinger & Swineford, 1937).

In 1957, Schmid and Leiman created their orthogonalization procedure. The SL procedure represents an elegant transformation of the higher-order model and is considered an approximate bifactor model (Reise, 2012). The bifactor conceptualization of intelligence essentially lay dormant until 1993 when John Carroll created his magnum opus, *Human Cognitive Abilities,* where he re-analyzed approximately 457 datasets going back to the 1920s. This creation ostensibly served as a bulwark against those who disavowed the importance and even existence of *g*. The creation of the SL transformation also raised awareness of the bifactor model as a tenable model for contemporary intelligence tests, but could only be produced secondarily from standard EFA results.

In 2011, Jennrich and Bentler created a true exploratory version of the bifactor model via analytic rotation (for further application and simulation of its use on real world data see Dombrowski et al., 2021). Prior to that time researchers who wanted an exploratory bifactor modeling approach had to utilize the Schmid-Leiman (1957) procedure. At present, both the higher-order model and the bifactor model may be used to investigate the structure of tests of cognitive ability though application of the former to other psychological measures is controversial as the theoretical justification for a general factor of say personality is less well-developed (Bonifay et al., 2017). Despite debate (see Decker et al. [2021] and a response via Dombrowski et al. [2021]) about which model reflects the true reality of the structure of tests of cognitive abilities and the nature of intelligence, there have been few studies that have empirically tested this issue (see Dombrowski, McGill, & Morgan [2021] for a Monte Carlo simulation of the structure of all major intelligence tests) so the issue remains yet unresolved.

[1] It is acknowledged that the Schmid-Leiman (1957) orthogonalization procedure is predicated upon the higher-order model transforming it into an approximate bifactor model where variance is apportioned to the general factor and group factors. This is different than Keith and Reynold's (2018) approach which apportions variance and preserves higher-order structure.

References

Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher- order models. *Journal of Intelligence, 3*, 121-136. https://doi.org/10.3390/jintelligence3040121

Beaujean, A. A., & Benson, N. F. (2019). The one and the many: Enduring legacies of Spearman and Thurstone on intelligence test score interpretation. *Applied Measurement in Education, 32*(3), 198-215. https://doi.org/10.1080/08957347.2019.1619560

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science, 5*(1), 184-186. https://doi.org/10.1177/2167702616657069

Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advances* (pp. 247-271). Hogrefe.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Decker, S. L., Bridges, R. M., Luedke, J. C., & Eason, M. J. (2021). Dimensional evaluation of cognitive measures: Methodological confounds and theoretical concerns. *Journal of Psychoeducational Assessment, 39*(1), 3-27. https://doi.org/10.1177/073428292094087

Dombrowski, S. C., Beaujean, A. A., Schneider, J. W. & McGill, R. J. & Benson, N. (2019). Using exploratory bifactor analysis to understand the latent structure of multidimensional psychological measures: An applied example featuring the WISC-V. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(6), 847-860. https://doi.org/10.1080/10705511.2019.1622421

Dombrowski, S. C., McGill, R. J., Canivez, G. L., Watkins, M. W., & A. A. Beaujean (2021). Factor analysis and variance partitioning in intelligence research: Clarifying misconceptions. *Journal of Psychoeducational Assessment, 39*(1), 28-28. https://doi.org/10.1177/0734282920961952

Dombrowski, S. C., McGill, R. J. & Morgan, G. W. (2021). Monte Carlo modeling of contemporary intelligence test (IQ) factor structure: Implications for IQ assessment, interpretation and theory. *Assessment, 28*(3), 977-993. https://doi.org/10.1177/1073191119869828

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: *g* as superordinate or breadth factor? *Psychology Science Quarterly, 50*, 21-43.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika,* 2, 42-54. https://doi.org/10.1007/BF02287965

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika, 76*, 537-549. https://doi.org/10.1007/s11336-011-9218-4

Keith, T. Z., & Reynolds, M. R. (2018). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 853-900). Guilford Press.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53-61. https://doi.org/10.1007/BF02289209

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696. https://doi.org/10.1080/00273171.2012.715555

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53-61. https://doi.org/10.1007/BF02289209

Spearman, C. (1904). General intelligence, 'objectively' determined and measured. First *American Journal of Psychology, 15*, 201-293.

Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review, 38*(5), 406-427. https://doi.org/10.1037/h0069792

Figure A1

Posterior Predictive Checking Scatterplot

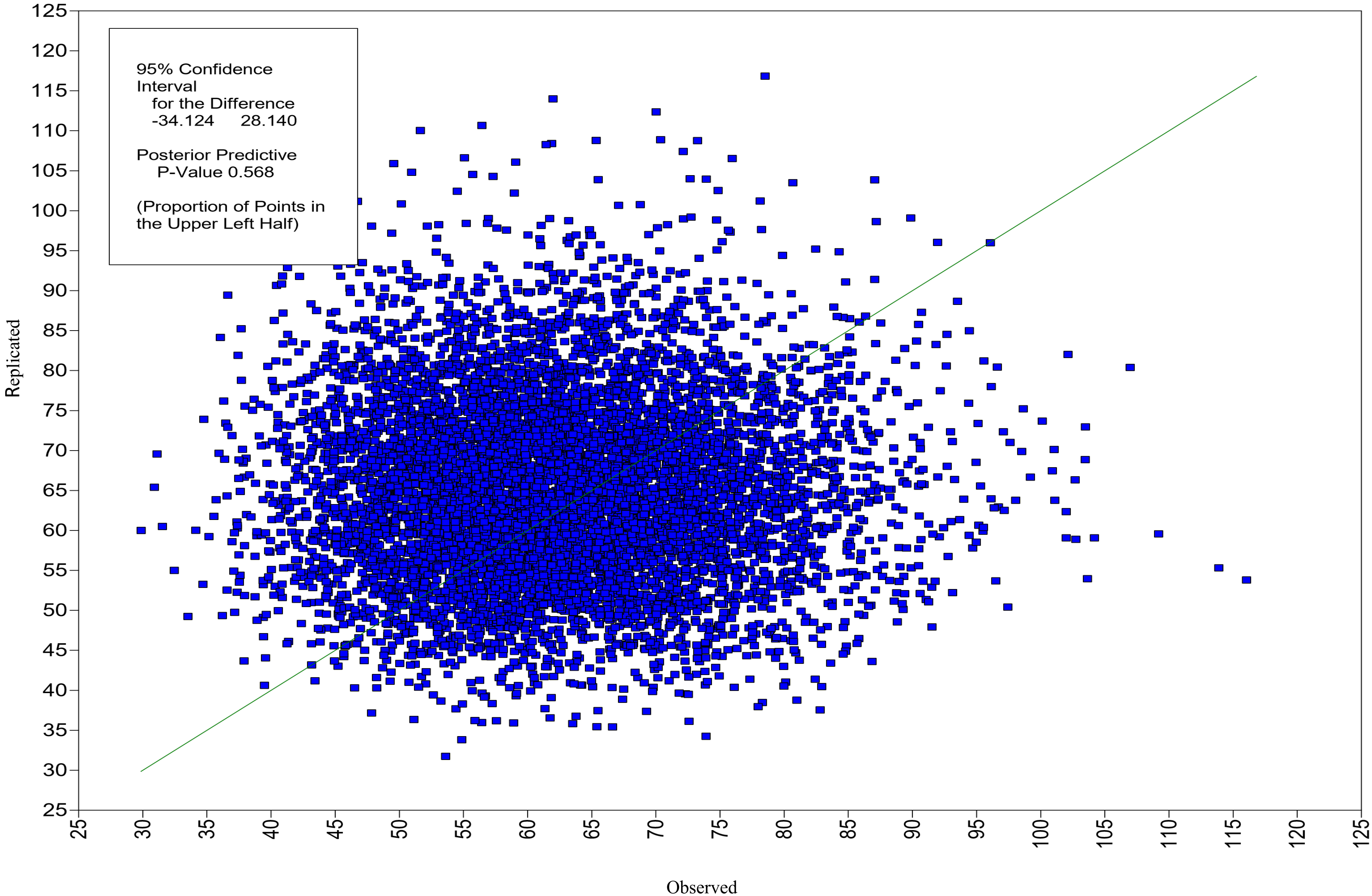


Figure A2

Posterior Predictive Checking Distribution Plot