PsychCorp

# Beta-4

## Beta–4
## Manual Supplement



ALWAYS LEARNING

PEARSON

PEARSON

*For inquiries and reordering:*
800.627.7271
www.PearsonClinical.com



PsychCorp

**NCS Pearson, Inc. 5601 Green Valley Drive Bloomington, MN 55437**

1.A Product Number 015800857X_1

# Introduction

The Beta–4 Manual Supplement provides additional information referred to in the Beta–4 Manual. This supplement is intended to augment the manual, allowing it to remain simple, uncluttered by detail, and portable.

## Test–Retest Reliability by Age Band

Table 1 presents the mean Beta–4 IQ scores and *SD*s for the first and second testing, the test–retest stability coefficient corrected for the normative sample's variability (Allen & Yen, 2002; Magnusson, 1967), the standard difference (i.e., effect size) between the first and second test sessions, and the standard error of measurement (*SEM*) for the Beta–4 IQ for each of two age bands (16–44 and 45+). The values for the overall sample appear in the Beta–4 Manual, Table 4.2.

**Table 1.** Test–Retest Reliability, by Age Band

**Ages 16–44**

| | First testing | | Second testing | | | | | |
|---|---|---|---|---|---|---|---|---|
| Score | Mean | *SD* | Mean | *SD* | $r_{12}$ | Corrected *r* | Standard difference | *SEM* |
| Beta–4 IQ | 97.6 | 13.8 | 104.5 | 15.6 | .89 | .91 | 0.47 | 4.50 |

**Ages 45–99+**

| | First testing | | Second testing | | | | | |
|---|---|---|---|---|---|---|---|---|
| Score | Mean | *SD* | Mean | *SD* | $r_{12}$ | Corrected *r* | Standard difference | *SEM* |
| Beta–4 IQ | 103.2 | 15.0 | 108.5 | 16.4 | .91 | .91 | 0.34 | 4.50 |

## *SEM* and Confidence Interval Formulas

The formulas for calculating the *SEM* and confidence intervals use various methods, including one used to develop the confidence intervals in the Beta–4 Manual, Table A.2. This section describes these methods.

The *SEM* is calculated with the formula:

$$SEM = SD\sqrt{1 - r_{xx}}$$

where *SD* is the standard deviation of the scale, and $r_{xx}$ is the reliability coefficient of the scale.

The *SEM* is used to calculate confidence intervals, or bands of scores around observed scores, in which true scores are likely to fall. Confidence intervals express test score precision and serve as reminders that measurement error is inherent in all test scores and that observed test scores are only estimates of true ability. Confidence intervals can be used to report an examinee's score as an interval that is likely to contain the examinee's true score.

There are multiple methods that can be used to calculate confidence intervals, and they all produce slightly different confidence interval values. All confidence interval calculation formulas take the *SEM* into account. Some local education agencies and other practice settings prefer to calculate confidence intervals in the most parsimonious manner, based only on the observed score and the *SEM*. Confidence intervals based on the observed score and the *SEM* are calculated by the following formula:

$$p\% \text{ Confidence Interval} = \text{Observed Score} \pm z_p(SEM)$$

where *p* is the confidence level (e.g., 90% or 95%), and $z_p$ is the *z* value associated with the confidence level, located in normal probability tables.

Depending on the preference of the practitioner or practice settings, the *SEM* for either the overall sample or for the examinee's age group can be used to calculate confidence intervals centered on the observed scores. The examples provided in this section utilize *SEM*s based on the overall sample. If an examinee obtained a Beta–4 IQ score of 108, the practitioner can be 95% confident that the examinee's true Beta–4 IQ score falls in the range of 99–117 (because the 95% confidence interval is 108 ±1.96 *SEM*, where the *SEM* is 4.74), and 90% confident that the examinee's true Beta–4 IQ score is in the range of 100–116 (108 ±1.65 *SEM*).

The confidence intervals for the Beta–4 IQ score ranges reported in the Beta–4 Manual, Table A.2, were calculated around the estimated true scores using the standard error of estimation (*SEE*), according to the method proposed by Dudek (1979) and Glutting, McDermott, and Stanley (1987). The estimated true score is obtained by the formula:

$$\text{Estimated True Score} = 100 + r_{xx}(X - 100)$$

where $X$ is the observed Beta–4 IQ score and $r_{xx}$ is the reliability coefficient of the Beta–4 IQ score. The *SEE* is derived by the formula proposed by Stanley (1971):

$$SEE = SD\left(r_{xx}\right)\sqrt{1 - r_{xx}}$$

where *SD* is the theoretical standard deviation of the IQ score, and $r_{xx}$ is the reliability coefficient of the IQ score.

This method centers the confidence interval on the estimated true score rather than on the observed score, and in turn, results in an asymmetrical confidence interval around the observed score. This asymmetry occurs because the estimated true score typically is closer to the mean of the scale than is the observed score. A confidence interval based on the estimated true score and the *SEE* is a correction for true-score regression toward the mean. For example, if an examinee obtains a Beta–4 IQ score of 113, the examinee's estimated true score will be 111.7, the 95% confidence interval of the examinee's true score will be 103–120 (because the 95% confidence interval is 111.7 ±1.96 *SEE*, where the *SEE* is 4.27), and the 90% confidence interval of the examinee's true score will be 105–119 (111.7 ± 1.65 *SEE*).

## Confirmatory Factor Analysis Interpretation

One of the most important outcomes of factor analysis is an understanding of the number and nature of the factors necessary to explain how the tests interrelate. For example, does the evidence suggest that there are one, two, or three distinct ability dimensions influencing performance on all the tests? Factor analysis is a tool for identifying the fewest factors that account for the data. Is there evidence that particular cognitive abilities (e.g., nonverbal ability) are among those dimensions? Examination of the content of the tests that are strongly related to a factor supports inferences about the nature of the ability represented by that factor. In this way, factor analysis links the measures in a test battery to existing measurement models.

The factor model proposed for a scale can be evaluated with a type of structural equation modeling (Jöreskog & Sörbom, 1993) known as confirmatory factor analysis. This technique is designed to evaluate a factor structure specified by researchers on the basis of theory and empirical research. The specific relations between observed variables (test scores) and latent variables (constructs such as nonverbal reasoning), and among the latent variables are specified in the model. In this sense, confirmatory factor analysis is different from exploratory factor analysis, because the tests are sorted into groups representing different factors in advance rather than generating the groups by applying a statistical algorithm to the data. The prespecified model is tested to determine if it provides a reasonably good, yet parsimonious, explanation of the actual correlations among the tests. The outcome of the analysis may suggest ways that the model could be improved to do a better job of explaining the data; however, such changes would be made only after careful consideration of whether they align with theory.

For several reasons, the technique of confirmatory factor analysis is preferred to exploratory factor analysis when an explicit theory of the factor structure is present or when there are competing models in the research literature (Schmitt, 2011; Stevens, 1996). The confirmatory method provides a numerical evaluation of the quality of a specified model, and it enables the researcher to compare the effectiveness of different models. Also, when the hypothesized ability dimensions are significantly correlated with one another, the confirmatory approach is more robust than exploratory factor analysis, whose results can be quite sensitive to small differences in the pattern of correlations.

## FIT STATISTICS

The effectiveness of a factor model is evaluated using goodness-of-fit statistics. Most such indicators are based on the chi-square statistic. The likelihood-ratio chi-square is a statistic used to test the null hypothesis that the matrix of subtest variances and covariances implied by the factor loadings and factor intercorrelations is the same as the observed variance-covariance matrix (Byrne, 2001; Schumacker & Lomax, 2004; Thompson, 2000). Therefore, if the model adequately fits the data, the null hypothesis is not rejected.

However, when the sample is large, chi-square tends to detect differences between the actual and implied variance-covariance matrices and to reject the null hypothesis even when the differences are small. For that reason, evaluation of model fit relies on other statistics that are not sample-size dependent (Byrne, 2001; Schumacker & Lomax, 2004; Thompson, 2000). Chi-square remains necessary, however, because of its utility in testing the statistical significance of a difference in fit between two models.

The Beta–4 analyses report additional model-fit statistics that are commonly used to determine the degree to which the observed and implied variance-covariance matrices are equivalent. The comparative fit index (CFI; Bentler, 1990) evaluates how well the specified model reproduces the observed variance-covariance matrix as compared to a model that presumes that all of the tests are uncorrelated. Hu and Bentler (1999) recommend a CFI value of .95 or higher as a criterion for good fit. The Tucker-Lewis non-normed fit index (TLI; Tucker & Lewis, 1973) measures the improved fit in a manner similar to the CFI, with an adjustment for the degrees of freedom in the model. Steiger's (1990) root mean squared error of approximation (RMSEA) is used to explore model fit relative to a population covariance matrix. The RMSEA is an estimate of how well the estimated variance-covariance matrix matches the population matrix, adjusting for the degrees of freedom in the fitted model. Browne and Cudeck (1993) suggest that a RMSEA value of .05 or less indicates a close model fit and that values of up to .08 represent adequate model fit with reasonable errors of approximation in the population. The Akaike information criterion (AIC; Akaike, 1974) and the Bayesian information criterion (BIC; Schwarz, 1978) are also included. When two models are compared on these statistics, smaller values suggest a better fit.

# References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. doi:10.1109/TAC.1974.1100705

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory.* Long Grove, IL: Waveland Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. doi:10.1037/0033-2909.107.2.238

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Beverly Hills, CA: Sage.

Byrne, B. M. (2001). Structural equation modeling: Perspectives on the present and the future. *International Journal of Testing, 1*(3-4), 327–334. doi:10.1080/15305058.2001.9669479

Dudek, F. J. (1979). The continuing misinterpretation of the standard error of measurement. *Psychological Bulletin, 86*(2), 335–337. doi:10.1037/0033-2909.86.2.335

Glutting, J. J., McDermott, P. A., & Stanley, J. C. (1987). Resolving differences among methods of establishing confidence limits for test scores. *Educational and Psychological Measurement, 47*(3), 607–614. doi:10.1177/001316448704700307

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. doi:10.1080/10705519909540118

Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8 [Computer software]. Chicago, IL: Scientific Software International, Inc.

Magnusson, D. (1967). *Test theory.* Reading, MA: Addison-Wesley.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment, 29*(4), 304–321. doi:10.1177/0734282911406653

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. Retrieved from http://www.jstor.org/stable/2958889

Stanley, J. C. (1971). Reliability. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 356–442). Washington, DC: American Council on Education.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173–180. doi:10.1207/s15327906mbr2502_4

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 261–283). Washington, DC: American Psychological Association.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1–10. doi:10.1007/BF02291170

# Appendix A. Special Group Criteria

Individuals were eligible to participate if they met all of the following criteria:

Primary language is English (with the exception of the Spanish-speaker and the English as a second language groups);

Able to communicate at a level commensurate with diagnosis, and not completely uncommunicative;

Normal hearing and vision (with aid);

Normal fine and gross motor ability (with the exception of mild motor impairment occurring in the intellectual disability group);

No physical conditions, illnesses, or impairments that could impact cognitive functioning or test performance (with the exception of conditions or impairments associated with a specific special group);

No diagnosis of a neurological condition (e.g., seizure disorder, epilepsy, encephalitis, brain surgery, brain tumor) other than the condition of interest or as allowed for a given special group;

No period of unconsciousness not related to surgery or greater than 20 minutes related to a medical condition;

No diagnosis of intellectual disability (with the exception of the intellectual disability group) or of a pervasive developmental disorder;

No diagnosis of a psychiatric disorder (e.g., psychotic disorders, mood disorders) other than that defined by the special group criteria;

Not currently admitted to a hospital, inpatient treatment, or psychiatric facility (with the exception of specified treatment settings that are pertinent to a given condition of interest [e.g., placements for correctional group or intellectual disability group]);

Not currently taking medication that might impact test performance, except as appropriate to treat condition of interest or associated conditions;

Meets the criteria for one special group only (with the exception of dual diagnoses of ADHD secondary to intellectual disability); and

Has not completed the Beta III or any other measure of cognitive ability in the 6 months prior to the testing date.

## SPECIFIC INCLUSION CRITERIA FOR SPECIAL GROUPS

### Correctional

Participation criteria included:

currently incarcerated in a correctional facility.

### Intellectual Disability

Participation criteria included:

meets *DSM-5*™ criteria for a current diagnosis of intellectual disability, mild or moderate severity;

OR

full scale score 2–4 *SD*s below the mean on a standardized, individually-administered measure of cognitive ability (e.g., IQ = 40–70).

## Attention-Deficit/Hyperactivity Disorder

Participation criteria included:

meets *DSM-5* criteria for a current diagnosis of ADHD;

has not taken ADHD medication for at least 24 hours prior to testing.

## Spanish Speakers

Participation criteria included:

endorsed Spanish as his or her primary language.

## English as a Second Language

Participation criteria included:

speaks and understands English well;

endorsed another language as his or her primary language.

## Specific Learning Disorder or Disability

Participation criteria included:

history of specific learning disorder, by self-report;

OR

history of having received school services for specific learning disability, by self-report.

# Appendix B. Beta–4 Examiners and Sites

## BETA–4 EXAMINERS

Dawn Hafling Magers
James P. Simonds
Christopher Sumner
Sarah Kennedy
Barbara Vines
Sylvia Sanchez
Patrick Dallas
Bonnie Campbell
Tracy Larson
Lisa Brockhuizen
Linda Guttman
Aubrey M. Johnson
JoAnn Salerno
Jamira Grana
Kevin Crippen
Yvette Lightbourn
Yael Horowitz
Gail Tonkovich
Loraine Alderman
Denise Fancher Zaldana
Jennifer Japhet
Amy Warren
Ann Roberts
Candace McKinley
Paige Beal
Sheelah Zink
Helen Stayna
Julie Ferguson
Lori Liebing
Heidi Benson Rodriguez
Shelly Stripling
Maleah Bufford
Terri Garmon
Wendy Mickley
Patricia Hollinger
James Wahrer
Ronni Laing
Cathy Abels
Lori A. Bailey
Rena Wechter
Ashley Holt
Elaina LeGros
Kalli Kronmiller
Marie Anne Hoffmann
Germaine Graham
Bobbi Smith
Crystal Joson
Patricia Sullivan
Valerie Caruso
Barbara Rhinehart
Kathryn Salter
Nancy Klein
Brittany Ubben
Maria Schreiber
Alejandro Mancha
Mikala Wilder
Misty Lay
Danita Lackey
Krysta Adams
Alanna Lipinski
Heather Giao
Patricia Guarini-Reyes
Amy Holloway
Jane Russell
Khristy Garrett
Brenda (BJ) Crossley
Cheryl O'Heir
Lisa Leon
Beth Miller
Jacqueline Johnson
Monica Galvez
Ana Magallanes
Christopher Zuccaro
Jacqueline Harrison
Bronwyn Guiles
Diana Iborg
Kelly Paolisso
Stefanie Nance
Joanna Meyer
Mariangela Cavagna
Melissa Thompson
Madeline Wesh
Darlene Gustavson
Kylie Walker
Melissa Ussery
Christine Kane
Rachel Burkhaulter
Reginald Mitchell
Sharna Wood
Barbara Shumard

## BETA–4 SITES

| | |
|---|---|
| LFG & Associates, LLC | Round Rock, TX |
| Tracy Larson Private Practice | Newark, NY |
| Dr. Loraine Alderman Private Practice | East Meadow, NY |
| Sandilands, Inc. | Detroit, MI |
| Therapy Group of Waco | Waco, TX |
| CD Fulkes Middle School | Round Rock, TX |
| Victor Child Care | Victor, NY |