

UTILIZING ITEM HISTORY FROM REVIEW DRAFTS

Item statistics are incorporated as elements of item history data provided in Review Drafts. The history data is compiled each time the item is administered as a part of a test scored by WRIB. When utilized, the statistics within the item history can greatly enhance effective item selection by indicating items best suited for the particular test at hand. Thus, by identifying items with maximum utility for your unique test requirements, overall test quality will be improved.

ITEM HISTORY DATA

An example of an item and its history is shown below:

SeqNum: 1340 Category: 659- Key: C
Author:
Comments: 659.223

1. The "first-line" supervisor is by definition a/an:
 - A. top executive.
 - B. highly efficient supervisor.
 - C. immediate supervisor.
 - D. immediate assistant to the top executive.

HISTORY:

<u>Loc</u>	<u>Class</u>	<u>Admin Date</u>	<u>NumCands</u>	<u>Difficulty</u>	<u>Discrim</u>	<u>Pbis</u>	<u>NumInSub</u>
OJ	0874	05/14/2003	112	75.00	0.500	0.500	50
SB	2450	05/07/2003	44	63.60	0.360	0.280	29
HA	4241	10/29/2002	19	63.20	0.670	0.580	35
SB	7850	11/27/2001	271	81.90	0.270	0.320	30
VV	2554	04/28/2001	27	92.60	0.130	0.310	30

The item history contains the following:

1. Loc: identifies which WRIB member (Location) used the item.
2. Class: the four-digit job classification code ("WRIB Code") as described in the discussion of Review Drafts
3. Admin Date: the administration date of the test in which the item was included.
4. NumCands: the number of examinees (candidates) in this administration.

5. Difficulty: the percentage of examinees who correctly answered the item.
6. Discrim: the discrimination index, which indicates the difference in proportions of correct response between those with the highest and lowest scores for the test section. This number may range from -1.00 to 1.00.
7. PBis: the correlation index (a point-bi-serial correlation) which indicates the relationship of performance on the item with overall performance on the test section. This number may range from -1.00 to 1.00 where negative values indicate a negative relationship between item and test performance. This is frequently used as the measure of item discrimination. In fact, it will have values comparable to the discrimination index when item difficulties are in the moderate-to-high range (approximately .75). However, when difficulty is extremely high or low, the correlation index can be artificially high. Therefore, if the discrimination index and point-bi-serial index do not agree, one should base item selection upon the discrimination index.
8. NumInSub: the number of items in the test section.

USING ITEM HISTORY TO SELECT ITEMS

Of the data provided, discrimination is the single most important indicator of an item's usefulness in the context in which it was used. To be considered effective, the discrimination generally should be greater than positive 0.20. Discrimination values should be higher (at least 0.30) for more "concrete" or precise categories, such as mathematics. For some of the "softer" areas, such as interpersonal skills or public contact, values in the range of .15 to .20 are often the best that are available.

The following table may serve as a general guide for item selection based on discrimination index values:

<u>Discrimination</u>	<u>Item Effectiveness in Context Used</u>
0.40+	Very Good
0.30-0.39	Reasonably Good
0.20-0.29	Marginal
Below 0.19	Poor

The key point is that the context in which an item is used has a significant impact on item effectiveness. Consistently low discrimination values indicate a poor item. However, a low discrimination value for an individual administration may indicate that the item was ineffective in the context in which it was used. Items are not effective in all situations, so to select the most appropriate items, you should incorporate information provided by the other elements of the item history with the discrimination indices.

The first element to consider is the WRIB code. A breakdown of the code enables you to determine factors that may affect item usefulness. For example, items effective when administered to clerical applicants may prove ineffective when administered to supervisory applicants, or items that work for journey level classifications may not work as well at the trainee level. Therefore, it is important to determine the applicant groups to which the item has been effectively administered.

The number of examinees should also be taken into consideration. A large number of examinees will produce more reliable data, whereas a small sample size may produce less reliable data. Therefore, you should take note of the sample size when interpreting item analysis data.

Item difficulty plays a significant role in the evaluation of discrimination values. If an item is too easy or difficult for the applicant group tested, the discrimination will be distorted. An item of extremely high or low difficulty cannot possibly have a high discrimination index.

The number of items in a test section (NumInSub) also affects discrimination values since discrimination relates item performance to performance on the test section. A small value will not provide sufficient stability of data for effective evaluation. A test section should contain at least 20 items, and preferably 25-30 items, to produce useful item analysis data. Thus, you should consider the number of items in a section, both when interpreting discrimination values from prior use as well as when developing tests.

USING SOURCE CODE TO SELECT JURISDICTION SPECIFIC ITEMS

A few categories have been designed to accommodate items that are jurisdiction specific, e.g., Law Enforcement Procedures - Department Procedures (621). Members who have items specific to their jurisdiction in one of these categories should, when needed, request those items by source code. This should be indicated on the Review Draft Request Form under "any other considerations." For example, "select items in category 621 by source code SB." Each jurisdiction that is a member of the item bank has a designated source code. Appendix J, p.55, contains the WRIB Source Code listings in alphabetic order by source code and jurisdiction.