Meta-analysis of OA and OAA manual determinations

Within the last month Stevan Harnad and his group, and I with my associates, have reported several manual measurements of OA (and sometimes OAA). The intent has been to evaluate the accuracy of Chawki Hajjem's robot program, which has been extensively used by Harnad's group.

Our group has now prepared an overall meta-analysis of the manual results. We are able to combine the results: we all were careful to examine the same sample base using identical protocols, for both the counting and the analysis, and measured a between-groups inter-rater agreement of 85%. Details of the measurements are being reported elsewhere.

We therefore have a combined sample of
1198 articles in biology and sociology,
559 of which the robot had identified as OA.
Of these 559 articles , only 224 actually were OA (37%)
The discriminability index, a commonly used figure of merit, was only 0.96667. For the full tables and graphs, see below.

The interpretation of these results will be discussed separately.

Dr. David Goodman
Associate Professor
Palmer School of Library and Information Science
Long Island University
dgoodman@liu.edu

Kristin Antelman
North Carolina State University Libraries
<kristin_antelman@ncsu.edu>

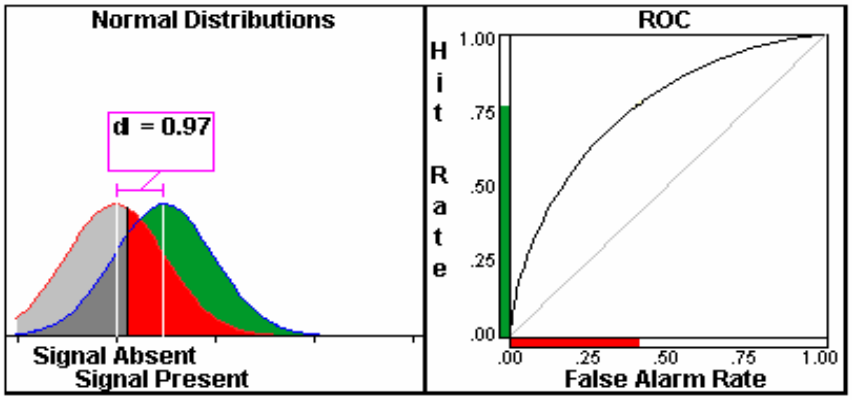
Nisa Bakkalbasi
Yale University Library
nisa.bakkalbasi@yale.edu

Jan. 3, 2005

Decision Table				
Manual Detection				
		OA	non-OA	TOTAL
Algorithm Detection	OA	224	375	599
	non-OA	66	533	599
	TOTAL	290	908	1198
	Probability	Z-Score		
Hit rate	0.77241	0.74682		
False alarm rate	0.41300	-0.21985		
d'=z(H)-z(F)	0.96667			
ß=e -[(z(H)^2-z(F)^2)/2]	0.77515			

Sample size:

Biology 2002	544
Sociology 2000	354
Sociology 2002	100
Chawki Hajjem	200
Total	**1,198**



Hits
Correct Rejection
False alarm
Miss

ROC: Receiver Operating Characteristic
Hit rate: 0.77
False alarm rate: 0.41