

## Meta-analysis of OA and OAA Manual Determinations

Dr. David Goodman  
Palmer School of Library and Information Science  
Long Island University [dgoodman@liu.edu](mailto:dgoodman@liu.edu)

Jan. 3, 2005

Kristin Antelman  
North Carolina State University Libraries [kristin\\_antelman@ncsu.edu](mailto:kristin_antelman@ncsu.edu)

Nisa Bakkalbasi  
Yale University Library [nisa.bakkalbasi@yale.edu](mailto:nisa.bakkalbasi@yale.edu)

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.

Decision Table				
Manual Detection				
Algorithm Detection		OA	non-OA	TOTAL
	OA		224	375
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			
$\beta=e^{-[(z(H)^2-z(F)^2)/2]}$	0.77515			

### Sample size:

Biology 2002	544
Sociology 2000	354
Sociology 2002	100
Chawki Hajjem	200
<b>Total</b>	<b>1,198</b>

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

