Empirical Assessment of Effect of Publication Bias on a Meta-Analysis of Validity Studies on University Matriculation Examinations in Nigeria

ADEYEMO Emily, Oluseyi
Faculty of Education Obafemi Awolowo University
Ile-Ife, Nigeria
E-mail:seyiadeyemo2007@yahoo.com

Received: January 30, 2012 Accepted: February 13, 2012 Published: June 15, 2012
doi:10.5430/wje.v2n3p57 URL: http://dx.doi.org/10.5430/wje.v2n3p57

Abstract

This study examined the impact of publication bias on a meta-analysis of empirical studies on validity of University Matriculation Examinations in Nigeria with a view to determine the level of difference between published and unpublished articles. Specifically, the design was an ex-post facto, a causal comparative design. The sample size consisted of 30 empirical studies selected on the basis of empirical status and relevance. The characteristics of these studies were recorded, coded and analyzed. The results revealed that the impact of publication bias was minimal on the meta-analysis study since the two versions i.e. published and unpublished articles yielded essentially similar estimates of effect sizes. (Mean Fisher Zr =0.393, Weighted Fisher WZr =0.398 with associated r =0.375 for both). Even though the use of non-representative proportion of significant studies in the two positive directions led to a non-representative set of studies, published articles had a higher effect size than the unpublished articles. (Zr = 0.460 for published, while Zr =0.279 for unpublished). The effect size of published articles was significantly different from unpublished articles ($\chi 2$=5.42 $p < 0.05$ sig). This was an indication that studies with more significant results were the published. On the overall, there was no significant difference in the level of significance for both published and unpublished articles ($\chi 2$ = 0.0031 $p > 0.05$). Hence, publication bias did not affect the results of this meta-analysis.

Keywords: Publication bias, Meta-analysis, Validity, Matriculation examinations, Effect size

1. Introduction

Publication bias is the tendency of researchers to handle the report of experimental results that are positive differently from results that are negative (i.e. report that shows a significant finding differently from results that support the null hypothesis) or inconclusive, leading to bias in the overall report. Such bias occur despite the fact that studies with significant results do not appear to be superior to studies with null significant results with respect to quality of design. Publication bias occurs when publication of research results depend on their nature and direction.

In meta-analysis, publication bias is a particularly thorny issue because meta-analysis has been put forward as providing a more accurate appraisal of a research than the traditional narrative review. If the sample of studies retrieved for review is biased, then the validity of the results of a meta-analysis review, no matter how systematic and thorough it might be is threatened.

Publication bias is a widespread problem that may seriously distort attempts to ensure the effect under investigation since the study is reviewed to determine features of the design and execution of both single studies by meta analysis. Also, the role the authors, journal editors and reviewer play in selecting studies for publication could lead to bias. The design of the meta-analysis itself and the studies included are also found to be important among a number of factors that result to publication bias. Study characteristics peculiar to each of the study can cause differences in the results, which can also lead to publication bias in the meta-analysis. There are various factors that can influence an author's decision to submit results for publications, journal editors and reviewers are crucial in deciding which studies to publish.

Publication selection effects arise in meta-analysis when the magnitude estimate are observed in only a subset of the studies that were actually conducted. From the findings of Larry and Hedges (1982) it was discovered that in research selection process much of the selection occurs because researchers, reviewers and editors review the result of studies as more conclusive when they are more highly statistically important i.e. they follow a model of selection process that depends on effect magnitude via the p-value or significant level. This according to Sterne, Gavagha, & Egger (2000)

may likely affect small studies which also tend to be of lower methodological quality. This may lead to small study effects where smaller studies in a meta-analysis show larger treatment. Small study-effect according to Sterne et al (2000) may also arise because of between trial heterogeneity.

Publication selection bias is difficult to document because of the nature of the phenomenon. Information about the missing estimate is needed. Among this is the failure to report the non-significant results, which occur when the mean difference are not statistically significant. Authors sometime, fail to report either the test statistics or descriptive statistics but simply report no significant difference. The tendency to report only the statistical analysis in detailed when significant results are obtained is part of what can lead to publication bias in the meta-analysis research. It has been found that statistically significant results are three times more likely to be published than papers affirming a null result (Dickensunk and Chalnens (1982). It has also been found that the most common reason for non-publication is the failure of an investigator to submit underlining researchers role in publication bias phenomenon (Eastrbook, Berlin & Gopalan , 1991).

In an attempt to reduce the problem of publication bias some prominent journals require registration of a trial before it commences so that unfavorable results are most withheld from publication. In addition, attempts to identify unpublished studies have proved difficult and unsatisfactory, but there is also the caution in the use of small and non randomized trial because of high susceptibility to error and bias. There is also the issue of the "file drawer problem" which according to Rosenthal (1979) is the tendency for negative or inconclusive result to remain unpublished by authors. This occur when many studies in a given area of research are conducted but never reported and those not reported may in average recorded different results from those that are reported. An extreme Scenario is that a given null hypothesis may in fact be published and show a statistically significant result while the highest percentage rejected and languish in the researchers file drawer. The percentage of the studies lost in the file drawer according to Jeffrey and Scargle (2000) can result in a significant bias. Publication bias has much effect on meta-analysis as studies may not be truly representative of all valid studies undertaken and hence may distort the results of the meta-analysis and systematic review of large number of studies.

To minimize the effect, researchers need to perform thorough search for unpublished studies. Since test for publication bias rely on the underlying theory that small studies with small sample size (i.e. large variance) would be more prone to publication bias ,while large scale studies would be less likely to escape public knowledge and more likely to be published regardless of significant finding. When studies conducted in a field are smaller, the effect sizes are also small. The remedy to this bias among others is to enhance on better powered studies with high quality research standard.

Meta-analysis is a quantitative methods for synthesizing results from individual studies to estimate an overall effect. Since its introduction, it has become an important tool in providing quantitative overviews of areas where many such individual studies have been carried out. One of the major defect to using meta-analysis has been the possibility of publication bias. The study for meta-analysis are usually chosen through a literature review. In this situation an inherent selection bias may arise since for example studies may tend to be published more readily if they are statistically significant or if they are more interesting in terms of the impact of the outcomes or studies may be suppressed by vested interest and hence to the loss of the authors. These can ultimately influence the results.

Bias also arises due to the use of a non-representative proportion of significant results or studies in a positive direction. This could lead to a non-representative set of studies in the meta-analysis. Meta-analysis which is the quantitative synthesis of results from multiple studies of a single scientific phenomenon is being increasingly used in many fields. Rosenthal(1991). The preferential publication of studies with the clear-cut or compelling results by journal may introduce an important bias in meta- analysis. If meta-analysis that has been put forward to provide adequate and impartial result about findings of previous researchers would be affected by publication bias, then, the validity of the result no matter how systematic or thorough is therefore threatened or if possible invalid

It is important therefore to address bias not only to ensure the integrity of the individual meta-analysis but also to ensure the integrity of research. When a meta-analysis that was published ignored the potential for bias but later discovered that the results was incorrect, the perceptions would be fostered among researchers that meta-analysis cannot be trusted. Hence encouraging the prevention and the assessment strengthens its use and usefulness. This study therefore examined the impact of publication bias on a meta-analysis of empirical studies on the validity of University Matriculation Examination in Nigeria. It identified the extent to which the results of the meta-analysis study had been threatened by publication bias and assessed the degree to which publication bias impacted the results of the meta-analysis.

2. Research Questions

(i) What is the extent to which study characteristics peculiar to the selected validity studies contribute to variance in the strength of previously reported validity studies on Matriculation Examinations in Nigeria ?
(ii) To what extent does publication bias impact the mean effect size in a random effect meta-analysis?
(iii) What is the effect size of both published and unpublished articles?

3. Research Hypothesis

(1) The effect size of journal articles on validity of Matriculation Examinations is not significantly different from non-journal articles.
(2) There is no significant difference in the level of significance for both published and unpublished articles on the validity of Matriculation Examinations in Nigeria.

4. Methodology

The research design was an ex-post facto. The sample size was made up of 30 validity studies on Matriculation Examinations in Nigeria purposively selected on the basis of relevance and empirical status from the population of studies conducted on Matriculation Examinations in Nigeria. Among the selected studies were 14 published and 16 un-published articles The study characteristics peculiar to the selected studies were recorded and coded. The qualitative results of each study were recorded and converted to common Effect size 'r'. The data collected from the empirical studies were analyzed using the Hunter and Schmidts (1990) psychometric meta-analysis method. Statistical techniques such as descriptive statistics like means and standard deviation as well as chi-squared test for diffused and focused tests were used to study the association of the coded characteristics across selected studies and these were analyzed using appropriate statistical methods.

5. Results

5.1 Research Questions 1: What is the extent to which study characteristics peculiar to the studies contribute to the large amount of variance in the strengths of previously reported validity studies on University Matriculation Examinations

The twenty characteristics features coded for this study were grouped into nine categories. These were classified as characteristics under the researcher's control. These were characteristics the researcher had under his control that he could influence upon. The primary researchers could manipulate these study characteristics at his own free will.

In research works, most of the decisions were based on theory and convenience. In some cases, some decisions were made by primary researchers not because they were appropriate theoretically but because they were the most convenient in terms of time factor; finance and proximity of location of where the study was to be carried out. Any of the decisions based on these selections could influence or alter the results.

Table 1 showed all the coding for the study characteristics of the 30 empirical studies. These codes were added together and recorded. The summation were used as 'weights'. The 'weights' were ascribed as independent variables because to a certain extent, the result recorded by primary researchers were influenced by these study characteristics.

<Table 1 about here>

Studies with bigger Weights(i.e. $W \geq 35$) has a major representations in the calculations of r. The studies whose weights were below 20 ($W < 20$) has no representations in the calculations of r, while W between 35 and 20 has a fair representation in the calculations of r. The sum of the indices gave the weights of each of the study with respect to what the primary researcher considered to be the most acceptable validity indices. The highest weight assigned to a study was 38; while the lowest weight was 29. The maximum weight for any study was 47 based on the coded characteristics. The bigger the weights, the more representative were the indices on which the primary researcher based his calculation of coefficient r. Studies 8 and 23 and 27 with weights 27 and 38 were biased by characteristics under researchers control, since the study with lowest weight has the highest r.

5.2 Research Question 2: To what extent does publication bias impact the mean effect size in a random effect meta-analysis

<Table 2 about here>

In substituting the various weights (W) and Fisher($\overline{Zr}$) from table into the equation above

Weighted Fisher (WZr) = 0.398 with associated r = 0.375 and the Mean Fisher($\overline{Zr}$) = 0.393 with the associated r = 0.0375. $\overline{WZr} = \overline{Zr}$ = 0.375 also.

The implication of this was that weighting by characteristics under the researcher's control did not contribute to any large amount of variance in the strength of previously reported validity of UME. This implied that characteristics under researcher's control peculiar to this study did not affect the obtained 'r's for the 30 validity studies. This was an indication that characteristics under the researcher's control did not have any impact on the mean effect size in a random effect meta-analysis. Publication bias did not have any impact on the result of the meta-analysis study..

5.3 Research Question 3 What is the effect size of both published and unpublished articles.

<Table 3 about here>

Effect size of published article is 0.460 and the effect size of unpublished article is 0.288 .This implies that published articles has a high validity while the unpublished article had the lower validity. The published article had a higher representation in the calculation of r and the unpublished had the fairly low representation.

<Tables 4 & 5 about here>

From the descriptive tables, 16 studies were unpublished and 14 were published. The unpublished non-journal articles included PhD and Master's theses. The Mean effect for unpublished articles was 0.460 while the mean effect for published articles was 0.288, therefore testing for the hypothesis.

Hypothesis 1: The effect size of published journal articles on validity of UME is not significantly different from unpublished non-journal articles.

<Table 6 about here>

In substituting values for $\left(Zrj - \bar{Z}\right)$ and (N – 3) in the above expression.

$$\bar{zr} = \frac{(237.4x0.460) + (804.56x288)}{1035.96}$$

Zr = $0.3\overline{29}$

$\chi^2 = 237.4\ (0.460\text{-}0.329)^2 + 801.56\ (0.288\text{-}0.3299)^2$

χ2calculated = 5.421 with df = 1 and p = 0.05. (P < 0.05 significant). The observed (calculated) χ2 is greater than the critical table value. The null hypothesis is rejected. The effect size of published articles on validity of UME is significantly different from unpublished articles. Although there were fewer journal articles than unpublished yet there was a sizeable decline in magnitude of correlation coefficient 'r'. The published 'r' = 0.429 while the unpublished 'r' = 0.279, the difference of 0.150 in their effect sizes could be as a result of publication bias created from any of the variable. Furthermore, studies with more significant results were the published articles.

Hypothesis 2: There is no significant difference in the level of significance for both published and unpublished articles on the validity of matriculation examinations in Nigeria

<Table 7 about here>

The Mean effect M1 for unpublished articles was 0.460 while M2 for published articles was 0.288. Although ,there were fewer published journal articles than unpublished, yet there was sizeable decline in the magnitude of their mean effect going from published studies to unpublished. Their standard deviation was 0.219 and the mean difference was 0.174.

Therefore testing for the level of significance of both published and unpublished articles, the result of chi-square (χ2) is as presented in Table.8

<Table 8 about here>

χ^2 table at p = 0.05 df 1 is greater than the calculated value. Hence, there is no significant difference in the level of significance for both published and unpublished articles on validity of UME.

6. Discussion

It is well known that there may be a non-representative proportion of significant studies in the specific literature. Moreover studies which are of poor quality may be differentially published. A high quality paper even if it does not exhibit statistical significance may well be accepted where a low quality and insignificant result will fear less well. One would expect that studies with the major characteristics of r would have the highest effect size and also vice versa The study with the highest characteristics of r(i.e Weighted 38) was not the one with the highest value of r or Mean Fisher (**Zr)** because of differences in coding. The impact of publication bias was minimal since the two versions (published and unpublished articles) yielded essentially similar estimates of the effect size .Also the use of non-representative proportion of significant results or studies in the two positive directions led to a non-representative set of studies. thus establishing the works of Beggs (1994).

It is noteworthy that the published articles had higher effect sizes than the unpublished articles and the effect size of published articles was significantly different from unpublished. This was an indication that studies with more significant results were the published articles. This results tend to be consistent with the findings from a previous study by Sterne, Gavaghan, Egger and Epidemoil (2000). Also the effect size of the unpublished articles was low. This trend had been reported in previous studies on publication and related bias in meta- analysis that publication bias are more likely to affect small sample studies which also tend to be of lower methodological quality and this may lead to small study effects where the smaller studies in meta-analysis show larger treatment which may also arise because of between trial heterogeneity.

The results of the study also revealed that there was no significant difference in the level of significance for both published and unpublished articles on validity of matriculation examinations in Nigeria. In majority of cases, publication bias analysis will show that bias probably had little impact and did not affect the outcome of the study.

7. Conclusion

Publication biases were common within the sample of meta-analysis, but most of these cases did not affect the overall results. Nevertheless, in cases where publication bias analysis suggest that bias exist this can serve as a warning to researchers and practitioners to regard the initial results cautiously and to avoid an intervention or policy that could be useless or even harmful.

References

Cohen J (1977) "Statistical Power Analysis for Behavioural Sciences (revised Edition)" New York.

Copper H (1984) "The integrative research review a social science approach "Beverly Hills, Califonia: sage

Glass G.V & Stanley J.C. (1970) "Statistical methods in education and psychology" Englewood Cliffs, New Jersey: Prentices-Hall.

Hedges L.V. (1982) "Estimation of effect size from a series of independent experiments" psychological Bulletin, 92, pp 490-499.

Monsteller F.M. & Bush R.R. (1954) "*Selected quantitative techniques in* G. Lindsey (Ed) Handbook of Social Psychology Vol. 1. Theory and method pp. 287 – 334 Cambridge, Massachusetts Addison – Wesley.

Rosenthal R (1991) *Meta-analysis procedures for social research ed. Newbury Park CA. SAGE Production.*

Rosenthal R (1984) "*Meta-analytic Procedures for Social research "Beverly Hills California"* Sage Publication Inc.

Rosenthal R. & Rosnow R.L. (1984) "*Essential of behavioural research (*Methods and data analysis)" McGraw-Hill, inc.

Snedecor G.W.&Cochram G. W. (1967,1980) Statistical Methods (6^{th} & 7^{th} ed) Ames;Iowa State University Press.Dear, K and Begg C (1992), "An approach for assessing publication bias prime to performing a meta-analysis, Statistical Science 7, 237-245

Begg, C.B. and Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50, 1088–1101. http://dx.doi.org/10.2307/2533446

Cooper, H.M. (1979). Statistically combining independent studies: A meta-analysis of sex differences in conformity research. *Journal of Personality and Social Psychology*, 37, 131–146. http://dx.doi.org/10.1037/0022-3514.37.1.131

Cooper, H.M. (1998). *Synthesizing Research: A Guide for Literature Reviews*, 3rd edition. Thousand Oaks, CA: Sage

Dickern, K. Cha,. S, Chalmers T.C.(1987) Publication bias and Clinical trials

Easterbrook, P.J, Berlin, J.A. Gopalan, R. Mathews D.R.(1981) Publication bias in clinical Research. Lancet 337: 86772

Egger, M., Davey Smith, G. and Altman, D. G. (2000). *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Books

Egger, M., Davey Smith, G., Schneider, M. and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629

Jeffry D.S.(2000) Publication Bias :The file Drawer Problem in Scientific Influence. Journal of Scientific Exploration 14(2):94-106.

Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin*, 86, 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.

Song, F., Easterwood, A., Gilbody, S., Duley, L. and Sutton, A.J. (2000) Publication and other selection biases in systematic reviews. *Health Technology Assessment*, **4**(10)

Stern, J.A. Gavaghan D & Egger M (2000) Publication and related bias in meta-analysis .Power of statistical test and prevalence in the literature.Epidemoil,53(11).1119-29

Table 1. Coded Variables for the Selected Studies

S/N	Size of publication	Year code	Type of	Sample type coverage	Range restriction	Types of validity	Internal validity	Validity reporting error	Study qualities	Total weight
1	3	3	3	2	2	2	2	2	10	29
2	4	3	4	3	3	2	2	1	16	38
3	3	1	3	3	2	1	1	3	10	27
4	3	1	2	2	2	2	2	2	11	26
5	3	2	2	1	3	2	2	2	12	29
6	4	2	3	3	3	2	2	3	13	35
7	1	1	2	1	2	2	3	3	11	26
8	5	1	1	3	1	1	2	2	11	27
9	2	3	3	3	3	3	2	2	13	34
10	5	3	2	2	2	2	2	2	13	33
11	5	1	2	3	3	3	2	3	14	36
12	3	2	2	2	2	2	2	1	13	29
13	4	2	3	3	3	3	3	2	16	35
14	5	3	3	1	3	2	2	1	14	35
15	5	1	3	3	2	2	2	2	13	33
16	5	1	3	1	2	3	3	3	15	36
17	3	1	2	2	3	2	2	2	14	31
18	5	3	2	2	3	2	2	3	15	37
19	5	3	2	2	2	2	2	2	11	31
20	2	3	3	3	3	3	3	3	14	37
21	1	2	2	2	3	3	2	2	14	31
22	5	3	3	1	2	3	3	2	14	37
23	2	5	3	3	3	3	2	3	15	38
24	3	1	2	3	3	2	2	2	14	32
25	5	1	3	2	2	2	2	2	14	33
26	5	1	2	2	2	2	2	2	12	30
27	5	3	3	3	3	3	2	2	14	38
28	5	3	2	1	2	2	2	2	11	30
29	5	1	1	3	3	2	2	2	15	34
30	5	2	2	1	3	3	2	2	14	34

Table 2. Effect of Publication Bias

	R	**Zr**	**W**	**(W)(Zr)**
1	0.9	0.4118	29	11.9422
2	0.32	6	38	12.600
3	0.21	0.232	27	5.7564
4	0.47	0.5101	26	13.2626
5	0.04	0.04	29	1.16
6	0.28	0.2877	35	10.0695
7	0.09	0.0902	26	2.3452
8	0.86	1.1155	27	30.1185
9	0.18	0.182	34	6.188
10	0.09	0.0902	33	2.9766
11	0.57	0.6475	36	23.31
12	0.29	0.2986	29	8.6594
13	0.61	0.7089	39	27.6471
14	0.24	0.2448	35	8.568
15	0.04	0.04	33	1.32
16	0.70	0.8673	36	31.2228
17	0.21	0.2132	31	6.572
18	0.37	0.3884	37	14.370
19	0.30	0.3095	31	9.5945
20	0.48	0.533	37	19.721
21	0.31	0.3205	31	9.9355
22	0.42	0.4477	37	16.5649
23	0.43	0.4477	38	17.0126
24	0.12	0.1206	32	3.8592
25	0.36	0.3769	33	12.4377
26	0.62	0.725	30	21.75
27	0.36	0.3769	38	14.3222
28	0.03	0.03	30	0.9
29	0.36	0.3769	34	12.8146
30	0.78	1.0454	34	35.5436
	Mean Fisher	**0.393037**	**985**	**392.5449**
	Weighted FIsher	**0.398**		

$\sum(W) = 985$ where W represent study characteristics.

$$\overline{Zr} = \frac{\sum(weight)(Zr)}{\sum(weight)}$$

Table 3. Effect Sizes of Published and Unpublished Articles

(a) Published Articles

Study	**N**	**R**	**Zr**
8	1800	0.86	1.1155
11	100	0.57	0.6475
14	30	0.24	0.2448
15	40	0.04	0.04
16	120	0.70	0.8673
18	180	0.37	0.3884
19	54	0.30	0.3095
22	227	0.42	0.4477
25	78	0.36	0.3769
26	60	0.62	0.725
27	159	0.36	0.3769
28	212	0.03	0.03
29	42	0.36	0.3769
30	222	0.78	1.0454
n=14	**237.4**	**0.429**	**0.460**

(b) Unpublished Articles

Study	**N**	**r**	**Zr**
1	250	0.39	0.4118
2	558	0.32	0.3316
3	300	0.21	0.2132
4	121	0.47	0.5101
5	40	0.04	0.04
6	800	0.28	0.2877
7	30	0.04	0.0902
9	750	0.18	0.182
10	802	0.09	0.0902
12	123	0.29	0.2986
13	1379	0.61	0.7089
17	687	0.21	0.2132
20	60	0.48	0.533
21	180	0.31	0.3205
23	6462	0.43	0.4477
24	107	0.12	0.1206
n=16	**804.56**	**0.279**	**0.288**

ISSN 1925-0746 E-ISSN 1925-0754

Table 4. Descriptive Statistics for unpublished articles

Variables	N	Mean	SD
Sample size(N)	16	804.31	1555.886
N-3	13	801.31	1555.886
Effect size		0.325981	0.1970280
Level of significance		0.0369	0.02024

Table 5. Descriptive Statistics for both published articles

Variables	N	Mean	SD
Sample size(N)	14	278.86	479.580
N-3	11	275.86	479.580
Effect size		0.378443	0.2453216
Level of significance		0.0414	0.01703

Table 6. Effect Sizes of Published and Unpublished Articles

Publications	Variables				
	n	Mean of N	R	Zr	χ^2
Published	14	237.4	0.429	0.460	
Unpublished	16	804.56	0.279	0.288	5.42

$P < 0.05$ (significant)

- $\sum_{j=1}^{k} (Nj - 3)(Zrj - \bar{Z})^2$ is distributed for χ^2 with $k - 1$ df

Table 7. Categories/Status of Publication

Category	Unpublished	M1 unpublished
Effect size	16	0.460
Level of sig	16	0.0369

Table 8. Test of level of significance

Publications	Effect Size r	**Zr**	**N**	**N - 3**	**Level of Sig.**	χ^2
Published	0.429	0.460	7.423	234.4	0.041	0.0031
Unpublished	0.279	0.288	4.5680	801.56	0.0369	

$\chi^2_{calculated} = 0.0031$ $p > 0.05$ (not significant)