

**EXTERNAL EXAMINER SUGGESTED 2 PAPERS THAT LOOK AT # OF  
AUTHORS: HASLAM ET AL. (2008) AND JONES ET AL. (2008).**

**Number of Authors Predicts Influence on Evaluations of Journal  
Submissions**

A thesis submitted in fulfilment of the requirements for the Degree

of Master of Science in Psychology

in the University of Canterbury

by Likie Shawn Lim and

Andy Martens

University of Canterbury

2010

---

Address correspondence to: Andy Martens, University of Canterbury, Psychology

Department, Private Bag 4800, Christchurch 8140, New Zealand. E-mail:

andy.martens@canterbury.ac.nz. Phone: 64 (country code) 3 364 2987 ext.4947. Fax: 64

(country code) 3 364 2181.

## Table of Contents

<b>Acknowledgments .....</b>	<b>3</b>
<b>Abstract .....</b>	<b>4</b>
<b>Glossary .....</b>	<b>5</b>
<b>Introduction .....</b>	<b>6</b>
<b>Method.....</b>	<b>9</b>
<b>Results.....</b>	<b>3</b>
<b>Discussion .....</b>	<b>9</b>
Heading 1 .....	20
Heading 2 .....	23
Heading 3 .....	25
Heading 4.....	30
Conclusion .....	31
<b>References .....</b>	<b>32</b>

## List of Figures (Figures, Tables or Graphs)

<b>Figure 1: ANOVA Table .....</b>	<b>8</b>
<b>Figure 2: Within Subjects Factors .....</b>	<b>3</b>
<b>Figure 3: Between Subjects Factors .....</b>	<b>3</b>
<b>Figure 4: Multivariate Tests.....</b>	<b>4</b>
<b>Figure 5: Descriptive Statistics .....</b>	<b>4</b>
<b>Figure 6: Pairwise Comparisons.....</b>	<b>5</b>
<b>Figure 7: Descriptive Statistics 2 .....</b>	<b>5</b>
<b>Figure 8: Graph of Mean Favourability Ratings .....</b>	<b>6</b>

## Appendix

<b>Abstracts .....</b>	<b>I-VI</b>
<b>Information Sheet .....</b>	<b>A</b>
<b>Survey Consent Form.....</b>	<b>B</b>
<b>Pre-Evaluation Questionnaire .....</b>	<b>C</b>
<b>Debriefing Sheet .....</b>	<b>D</b>
<b>Verbal Debriefing Sheet .....</b>	<b>E</b>

# Abstract

180 students from the University of Canterbury were randomly assigned to reading and evaluating 4 counterbalanced abstracts under the cover story of a departmental journal submission procedure. This study tested whether the number of authors assigned to a journal submission is an influential factor on the acceptance rate of a submission regardless of the quality of the abstract. Also, it assessed whether the influence of a number of authors on the chance of acceptance interacts with the acceptance rate of the journal. In other words, the study investigated not only the extent to which number of authors influences acceptance regardless of quality, but how much of an influence this has for which kind of journals (in terms of the journal's acceptance rate). The study also measured how much individual personality variables such as guilt-proneness and tendency to adhere to descriptive norms influences a reviewer's willingness to accept a journal submission. Results found that number of authors had a significant effect on evaluation. Possible reasons and study limitations were discussed.

# Introduction

The development of scientific literature is dependent upon which submissions proceed into the final and decisive stage of publication. Progression of this sort requires acceptance by the reader or reviewer who is authorised to commend and merit certain articles. The criteria for acceptance of a journal submission would naturally vary across different journals and topics of academia. The normal procedure for article acceptance to journals usually involves reviewers appraising articles' quality. Which articles are published may also be influenced by factors unrelated to quality, however, and there is evidence of bias in which articles are accepted (Cherkashin, 2008). Factors such as the author eminence or level of recognition, and degree of popularity of subject matter might play their part in acceptance or rejection of a submission. We are examining an as yet untested possibility that the number of authors may bias a reviewer. Specifically, we think that more authors may bias reviewers/editors to give more favourable responses. To test this, we will ask our participants to review abstracts (as opposed to entire articles, for the sake of time).

There are a few reasons why we would expect more authors may lead people to be positively biased in the appraisal of a journal submission. First of all, guilt is a factor we believe may contribute to this kind of bias. Hypothetically, this is analogous to a situation

where one would find it harder to give a low score for a poem written by a group of five friends in a poetry writing competition over giving a low score for a poem written by one friend only. Basic principles of human nature would suggest that one would feel it harder to cause the “death” or “failure” of two people than one, three than two and so forth (Tangney and Dearing, 2003). The more authors seem to be involved in the production of an abstract, the more we expect people to feel worse if they reject it compared to rejecting an abstract produced by fewer persons.

Along similar lines, accepting an abstract written by more authors might make that the reviewer feel subconsciously that the authors would be grateful for his “approval”. Approving an abstract written by more authors may give the reviewer the illusion that they are helping more people. In a way, it is very similar to the reason of guilt, but this mechanism is manifested in a slightly different way – as opposed to avoidance of a greater guilt, they are actively pursuing a greater utility by “providing” success to more people.

The second main reason we could think of that could bias people favourably towards articles written by more people is that people follow descriptive norms. According to Perkins (2002), descriptive norms are perceptions of how other people are conducting behaviour and the “most common actions actually exhibited in a social group”. One previous investigation of our tendency to follow descriptive norms involved participants who were given the opportunity to litter into either a previously clean or a fully littered environment subsequent to witnessing a confederate who either littered into the

environment or did not (Cialdini, Reno, & Kallgren, 1990, experiment 1). In short, participants were prone to litter into an already littered environment than into a clean one because the environments created descriptive norms—one environment suggested that other people litter and the other situation suggested that other people do not litter. In a similar way, more authors may represent a descriptive norm and may lead reviewers to feel that more people support the paper's idea. More authors of an abstract would make it appear as if more people have agreed with the work, so reviewers might adhere to this descriptive norm by being more favourable towards the abstract. In other words, if people see more authors making a point, they may be slightly more inclined to want to agree with it too. This basic psychological mechanism could as a result, also turn into the false belief that the abstract is actually of better quality.

Another variable worth investigating is the interaction between the acceptance rate of a journal and the influence of number of authors. In such a case, the number of authors would be more or less influential for a journal with relatively low or high standards, in any order. Namely, it would mean that for journals with lower acceptance rates the ones with more authors should not be favoured over ones with less authors and have differing levels of acceptance, simply because the people in the position of reviewing these abstracts may pay more attention to the content and not enough attention to the number of authors to constitute any bias in making the final decision. This is the interaction which we are trying to find. Locating acceptance rates for individual journals or for specific disciplines can be difficult, yet is necessary information for promotion and tenure activities. Journals with lower abstract acceptance rates are frequently considered to be

more prestigious and more “commendable” or “praiseworthy”. This could signify the exception for the main hypothesis (given it is true for conditions which do not involve these “prestigious” journals). In other words, the phenomenon we are testing in this study may be negated in such a situation.

We will test our main hypothesis by providing participants with a cover story that they will review papers to be submitted to a new student publication/newspaper and that their votes are needed as the criteria for its acceptance. Initially, we will examine whether acceptance rate moderates the impact of number of authors by first telling participants that the journal has either a 50% acceptance rate or a 20% acceptance rate. That this, we will attempt to examine whether there is an effect of acceptance rate on the function of number of authors on abstract evaluation. Participants will then be asked to give their opinions about 3 abstracts, one with 1 author, one with 3 authors, and one with 6 authors. Each abstract has three versions, one with 1 author, one where with 3 authors and one with 6 authors. Additionally, in order to examine the potential role of guilt and adherence to norms in how favourably participants rate the article abstracts, we will also assess the individual variables of conscientiousness, agreeableness and guilt proneness. This will be done by attaching three separate questionnaires with each abstract distributed to the participants. These abstracts will be all counterbalanced for order to avoid confounding. Then, we will measure how favourably participants view the abstracts by asking them to rate how interesting each abstract is, how well written each abstract is, and whether they think it should be accepted or not.

Statistical analysis: 2 (acceptance rate: 20% vs. 50%) X 3 (number of authors: 1 vs. 3 vs. 6) ANOVA with number of authors as a within subjects factor (i.e., repeated measure).

	20% acceptance rate	50% acceptance rate
1 author		
3 authors		
6 authors		

If our hypothesis turns out to be correct, there will be an ascending level of acceptance rates from 1 author to 3 authors to 6 authors for journals with 50% acceptance rates. Also, this pattern will be significantly weaker for the abstracts with 20% acceptance rates.

# Method

## Participants.

150 people (25 per condition) were recruited from various locations around University of Canterbury. The data comprised of both male and female participants who were university students over the age of 17.

## Procedure.

Participants will be approached to participate in the study in various locations across the University of Canterbury campus. We presented a background that journal abstracts are to be submitted to a new University journal and the participants' votes were needed as the criteria for its acceptance as part of the publication process. However, this journal does not and will not exist and is merely a fabrication for the purpose of letting participants give answers as truthfully and naturally as possible. Participants were screened for age (participants are eligible if over 17 years of age) and 'offered' a chance to participate in the experiment masquerading as a journal cover story selection as an incentive.

Participants were informed that each abstract they read will be submitted and their votes will make up the criteria for acceptance into journal publication. Initially, the participants were recommended to evaluate each abstract based on how interesting it is and how

well-written it is, and that only a certain percentage of abstracts will eventually be accepted. Then, we described that the set of abstracts which they would soon read will be submitted to a journal with either a 20% or 50% acceptance rate, and that other participants will read sets of abstracts submitted to journals with one of these two acceptance rates, also. It would not make sense to tell them after they read the abstracts. Before each participant had finished reading all abstracts and answering all questions, we proceeded to control for certain individual difference variables for exploratory purposes. Namely, we attempted to measure the extent to which avoidance of guilt and adherence to descriptive norms influenced the decision-making process in the evaluation of abstracts. There were two questions asked assessing guilt-proneness, two assessing conscientiousness, and two assessing agreeableness. This stage was conducted prior to the participants reading of the abstracts, so as to avoid this manipulation distorting the participants' evaluation of them. Participants were then asked to read the abstract first before answering several questions about it to evaluate it. Participants read 3 abstracts. Each abstract has three versions, one matched up with 1 author, one matched up with 3 authors and one matched up with 6 authors. These abstracts were all counterbalanced. This means that each participant were seeing a totally different ordering of how many authors are attached to each version. For instance, one participant will read abstract #1 with 1 author while the next will see abstract #1 with 4 authors, while the next person will see abstract #1 with 6 authors. This technique not only reduces any unwanted ordering effects (people may score the first, second or third abstract they read higher or lower) but it gave us the chance to see if there is any effect of number of authors on reviewer response regardless of the quality of the abstracts (the quality of the abstracts

across each other will be similar enough to not be a considered a factor in the reviewers' decision making). After reading each abstract, participants were asked to rate the abstract on three questions:

- Do you vote to accept abstract or reject the abstract?
- How interesting was the abstract?
- How well written was the abstract?

Next, a manipulation check of the number of authors was commenced. This kind of test ensured that participants in different conditions of an experiment are experiencing different levels or conditions of the independent variable. In this particular study, we need to find out whether the participants were able to discriminate the differences in each of the conditions which were manipulated (differing number of authors). This was performed by asking each participant whether they remember how many authors each of the three abstracts they were given to read were written by. This is to test how conscious they were of the number of authors of each abstract. The more conscious the participants are, the more the experiment was successful in manipulating the variables and in turn achieve construct validity in the questionnaire. The second manipulation (of the acceptance rate) will be performed by presenting half of the abstracts to the participants as being submitted to a journal with a 20% acceptance rate and half of the abstracts as being submitted to a journal with a 50% acceptance rate. This will allow for analysis of whether there is a significant interaction between the acceptance rate of a journal and the influence of the number of authors. Demographic checks were also included. Gender and

age were also collected for exploratory purposes. Finally, participants were given a voucher as compensation and afterwards they were informed of the true purpose of the procedure and conditions which were manipulated to obtain this abstract.

# Results

Figure 2

## Within-Subjects Factors

Measure: MEASURE\_1

factor1	Dependent Variable
1	auth1comp
2	auth2comp
3	auth4comp

Figure 3

## Between-Subjects Factors

		Value Label	N
acceptrate_r	1.00	20%	87
	2.00	50%	93
order	1		34
	2		25
	3		33
	4		30
	5		42
	6		16

Figure 4

Multivariate Tests <sup>c</sup>						
Effect		Value	F	Hypothesis df	Error df	Sig.
factor1	Pillai's Trace	.047	4.080 <sup>a</sup>	2.000	167.000	.019
	Wilks' Lambda	.953	4.080 <sup>a</sup>	2.000	167.000	.019
	Hotelling's Trace	.049	4.080 <sup>a</sup>	2.000	167.000	.019
	Roy's Largest Root	.049	4.080 <sup>a</sup>	2.000	167.000	.019
factor1 * acceptrate_r	Pillai's Trace	.001	.116 <sup>a</sup>	2.000	167.000	.891
	Wilks' Lambda	.999	.116 <sup>a</sup>	2.000	167.000	.891
	Hotelling's Trace	.001	.116 <sup>a</sup>	2.000	167.000	.891
	Roy's Largest Root	.001	.116 <sup>a</sup>	2.000	167.000	.891
factor1 * order	Pillai's Trace	.020	.346	10.000	336.000	.968
	Wilks' Lambda	.980	.344 <sup>a</sup>	10.000	334.000	.968
	Hotelling's Trace	.021	.342	10.000	332.000	.969
	Roy's Largest Root	.014	.478 <sup>b</sup>	5.000	168.000	.792
factor1 * acceptrate_r * order	Pillai's Trace	.036	.617	10.000	336.000	.799
	Wilks' Lambda	.964	.614 <sup>a</sup>	10.000	334.000	.802
	Hotelling's Trace	.037	.610	10.000	332.000	.805
	Roy's Largest Root	.021	.692 <sup>b</sup>	5.000	168.000	.630

a. Exact statistic

b. The statistic is an upper bound on F that yields a lower bound on the significance level.

c.

Design: Intercept+acceptrate\_r+order+acceptrate\_r \* order  
Within Subjects Design: factor1

Figure 5

Descriptive Statistics				
	acceptrate_r	Mean	Std. Deviation	N
auth1comp	20%	4.557	1.3477	87
	50%	4.602	1.3545	93
	Total	4.581	1.3476	180
auth2comp	20%	4.713	1.2191	87
	50%	4.823	1.2061	93
	Total	4.769	1.2102	180
auth4comp	20%	5.040	1.2741	87
	50%	5.059	1.3650	93
	Total	5.050	1.3182	180

Figure 6

## Pairwise Comparisons

Measure: MEASURE\_1

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
					Lower Bound	Upper Bound
1	2	-.163	.118	.169	-.397	.070
	3	-.446*	.157	.005	-.755	-.136
2	1	.163	.118	.169	-.070	.397
	3	-.282*	.129	.030	-.537	-.027
3	1	.446*	.157	.005	.136	.755
	2	.282*	.129	.030	.027	.537

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

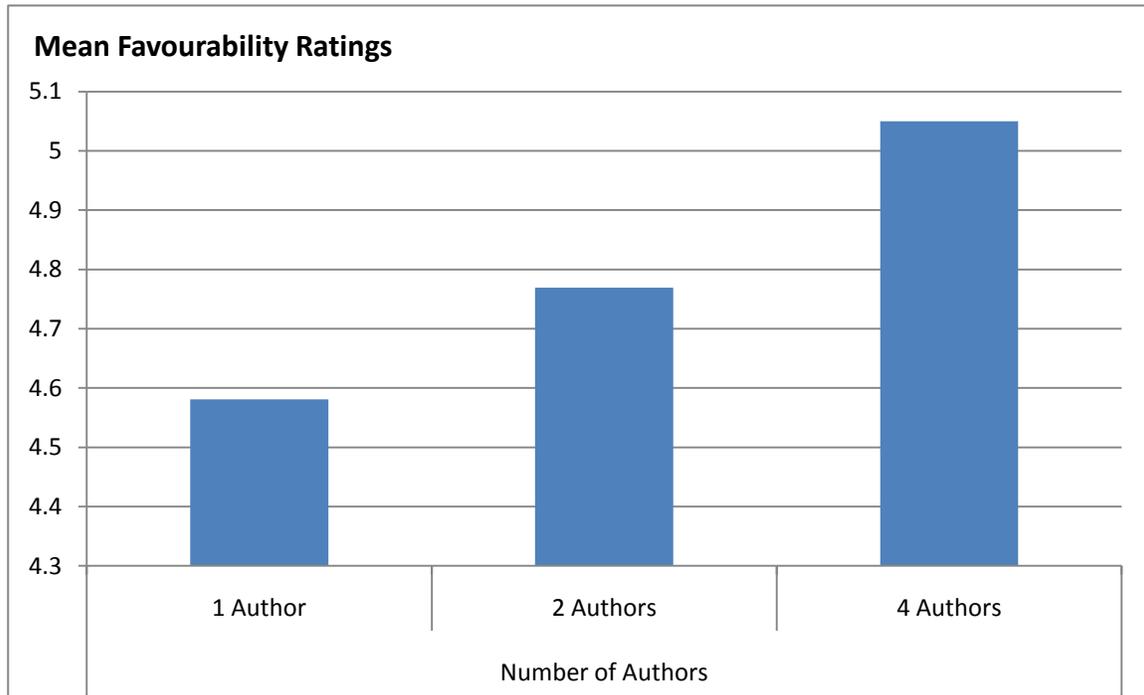
a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Figure 7

## Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
auth1comp	180	1.0	7.0	4.581	1.3476
auth2comp	180	1.0	7.0	4.769	1.2102
auth4comp	180	1.0	7.0	5.050	1.3182
Valid N (listwise)	180				

Figure 8



The above tables summarize the main findings reported.

Now, the main point of this study was to examine the different effects of the number of different authors on participant ratings of article summaries. There were also exploratory self-reported individual variables measured which bared no significant results, namely proneness to feel guilt and conscientiousness.

First, we examined participants' ratings (of how "interesting" the summary was to read and of how well written they thought the summary was) for each of the three separate abstracts. Specifically, for the dependent variable we average participant ratings. Submitting these average rating scores to a 2 (acceptance rate: 20% vs.50%) x 6 (order) x

3 (number of authors: 1 vs. 2 vs. 4) ANOVA revealed a main effect for number of authors,  $F(2, 167) = 4.08, p < 0.05$ . See **Figure 4**.

While there was no effect for order and acceptance rate and no interaction effects,  $ps > .18$  as reported in **Figure 4**), it was noted that number of ‘authors’ did have significant effects on the ratings. To more broadly examine this effect of number of authors, we investigated pair-wise comparisons using least significant difference tests. The next and perhaps most revealing result were the differences in the ratings given for the different numbers of authors. The mean for abstracts with 1 author was 4.581, the mean for abstracts with 2 authors was 4.769 and the mean for abstracts with 4 authors was 5.05. Ratings for abstracts written by four authors were significantly bigger than abstracts written by 2 and 1 authors ( $ps < 0.05$ ). However, ratings for abstracts written by 2 authors wasn’t significantly bigger than ones written by 1 ‘author’ ( $p = <0.169$ ). These findings are significant since they allow one to imply that abstracts appeared to have been written by more authors are viewed more positively by participant than abstracts written by fewer authors.

Neither the interaction between acceptance rate and number of authors ( $F(1,167) = 0.116, p = 0.891$ ), nor the interaction between order and number of authors, ( $F(1,167) = 0.346, p = 0.891$ ) nor the 3 way interaction between acceptance rate, order and number of authors ( $F(1,336) = 0.617, p = 0.799$ ) were significant. These are essentially positive results, because it implies that any interactions between the variables which were not hypothesized to directly influence the number of authors did not, giving more weight to the significant effect of number of authors.



---

# Discussion

The results pertained the only significant variable worth validating was the number of authors. In particular, ratings for abstracts written by “4 authors” were significantly higher than ratings for abstracts written by “2”.

The study investigated not only the extent to which the number of authors influence acceptance regardless of article quality, but how much of an influence this has for which kind of journals (in terms of the journal’s acceptance rate). These other variables of acceptance rate and order of the articles had no significant effect. Despite the fact that only one independent variable out of the numerous reviewed turned out to be significant, this was a generally positive result. As to whether the influence of a number of authors on the chance of acceptance interacts with the acceptance rate of the journal, there was no effect either. The singularity of the variable significance also untangles any potential confusion of unnecessary interaction between the variables. In turn, this simply led to a stronger argument for the number of authors variable. Not only do these results emphasize the actual significance of the presentation of number of authors in journal evaluation, it broadens the scope for future research into this one aspect for all kinds of literature and not just scientific or psychology in particular. There were various limitations and weaknesses in the study, however.

---

*Limitations Regarding the Number of Authors and Influence on Evaluations*

Firstly, limitations brought up in this study focused on the reliability of participants' judgments of article quality. First of all, the study is limited because the data collection only involved evaluation data from university-level students in a tertiary location with public access. And as these participants were not followed through for with identification checks or asked for a minimum qualification, there is no guarantee that the participants had the "ability" to read the articles thoroughly and accurately before evaluating them. Basically, were they qualified enough to judge which abstract/article is justifiably worthy of journal submission? Had this limitation been covered for, this data collection would have wholly fulfilled its role of the "cover story" of a psychology journal. Moreover, there was no "elite", handpicked exclusivity as to the characteristics of the authors. In other words, the "cover" authors were not given an introduction or background which is usually a significant factor in legitimate journal evaluation.

As well as these limitations of participant credibility, there were critical questions regarding individual characteristics such as guilt-proneness and conscientiousness which were explored but not probed to maximum potential due to the scarcity of reward for participation against the time in which such exploration would require.

Technically, it is also critical to note that the "articles" were not "articles" in full, but abstracts of non-existent "articles" reduces the significance of the study significantly – since it simulated a scenario where a university level student were approached to read

and evaluate scientific abstracts as opposed to qualified psychology professionals or academics evaluating whole articles for a psychology journal. These potential variables should perhaps be left for future discussion in the vicinity of this topic.

Another interesting issue that we could delve into would be the "informative" influence of the participants - which can be defined specifically in this study as the tendency for participants to conform to the general preferences of others regarding the "literary quality" and "interestingness" of the journal articles. In other words, the potential influence of whether participants possessed a sum of knowledge of each others' positions or opinions on the subject matter OR whether participants were given the freedom to communicate or argue their opinions for a less isolated, "socially" supported perspective before the actual evaluation takes place. This phenomena was most famously proven in Muiser Sherif's experiment in 1935 in which participants were asked to look at a dot of light 15 feet away in a darkened room and asked to estimate how far the light moved. Each participant was paired with two others and tended to give a common group estimate, which they maintained even when they were asked another time when they were alone – namely due to the ambiguity of the distance (Sherif, 1935).

Basically, perspective is restricted in the respect that the participants in this study were not given any information about other participants' preferences. The degree of ambiguity of the quality of the abstracts did not require participants to have this information for objective evaluation, yet such information could have tilted the direction of the study dramatically – especially the influence or the "presence" of the number of authors in consideration.

Furthermore, in most scientific journal evaluation, assessors would have a fair

knowledge of authors' credibility and reputation. However, the names of authors in this study were made up and assumed to be undergraduate and postgraduate psychology students according to the debriefing of the cover story prior to the evaluation. This adds to the limitations of the "informative" aspect of the study. Simply put, there was no knowledge presented as to which authors had higher reputation (or grades) than others. Given that this study only presented abstracts as to the whole articles, the mere "previews" to the authors' writing skill may not have been enough for a much more clarified participant opinion.

If one happens to write, edit or submit a scientific or political journal with a considerable amount of literary credibility already acquired under their belt, each point and paragraph would definitely add to the author's own public regard and high admissibility, albeit risking suffering endless comparisons to their own self-made measuring stick (Baxter, 2009). Nevertheless, this factor was ruled out in the present study which again strengthened the potential significance of the number of authors variable.

We refrained from asking whether the participants recalled whether they considered the authors' names (or number of authors) during their evaluation after the evaluation finished since that would unjustly rule out any implicit processing or sub-conscious effects they may not have been aware of at the time.

In fact, due to the scarcity and scheduling conflicts of psychology students in the University of Canterbury these major-specific prospect participants were not filtered down from the general university population.

### *How Can Future Studies Deal With Limitation*

First of all, future research more closely examining the participants before they proceed to undergo the evaluation by requiring identification of qualification would rectify the shortcoming of the reliability of participants' judgments. There should also be a check on which major the participants are currently enrolled in depending on the topic of the articles being evaluated in order to further minimize the range of participant backgrounds.

Participant reliability could also have been tested by giving participants the evaluation and then re-testing them again on the same articles a few days later on a follow-up study. Once the data is collected, participants who gave inconsistent assessments on the same abstracts would be flagged and their data also possibly scrapped for unreliability.

As for the limitation regarding the informative aspect of the study, the procedure of evaluation could have been altered by having multiple participants (preferably at least three or four) having a discussion after they are given time to read through each piece of literature before being asked to evaluate them individually. Not allowing the participants to communicate which positions they will actually vote for has been shown to have an effect on preference shifts in past studies as aforementioned. Particular studies in the past have purported that when arguments are shared in the absence of others' positions, strong shifts are found (Burnstein & Vinokur, 1975; Clark, Crockett, & Archer, 1971; St. Jean,

1970). Allowing each participant to have information on others' preferences would also cause a strong shift – yet would detract focus from the main hypothesis of the study.

As for the aforementioned lack of background information on the mask authors of the abstracts, there was plenty of room for minimum introductions for each author as an appendix to the evaluation. Participants would then be given a choice to refer to this appendix as a weighing factor in their evaluations. Since the evaluation was hierarchical in nature (numbers used to measure their favourability) this kind of information would have lead to clarity and satisfaction in the participants' choices.

Another interesting alternative to future potential replicas of this study would be to carry out a between-subjects design instead of (or as well as) a within-subjects design. To elaborate, instead of having the abstracts rotated with the different combination and number of authors; it would be interesting to study the effects if there was only one abstract presented for each participant. This may see a different result as the unseen interaction between the reading of the abstract and the presentation of authors' names may be reduced to a minimum.

Future research may also address an important limitation of the present study – that we assessed shortened versions of abstracts, and not full articles. Technically, the results of the present study may not hold much significance considering the evaluation was only of a part of the full “articles” and their publication of which the participants were “imagining” they were contributing to. In a real life situation, it would be unusual to not to let the qualified body of evaluators view more of the entirety of the article before the final evaluation is submitted. Future studies should focus on lengthening the abstracts

(or presenting the entire cover article) and duration of evaluation - which would also require higher rewards for mass participation.

Practically speaking, greater rewards for participation (e.g. gift vouchers and lottery tickets higher in value) would have allowed for a higher tolerance of more self-reported individual characteristics given as participant data and also time invested in contributing to the study. Although there was some kind of reward for the present study, it came across our mind that future studies dealing with similar subject matter would entitle a higher research scholarship for participant consent, given the positive results found. This would imply a higher number of participants, a self-report questionnaire more closely examining individual characteristics for exploratory purposes and reasoning for any effects, and longer abstracts (or articles) to be read. Basically, the main practical limitation to the study was reward, and time.

### *Reasons for Effect*

Despite the considerable agreement demonstrated here on prescriptive norms for the assessment of manuscripts and articles, and despite the demonstrated importance of peer-evaluation processes to the maintenance of our social enterprise (Brackbill & Korton, 1970; Garvey & Gottfredson, 1976), studies of peer-evaluation processes in psychology have offered a generally dismal picture. This study attempted to reroute this area of research by investigating the chosen variables of article evaluation mainly inspired by

rational ideas.

One of the prime reasons why people would prefer a higher number of authors would be the normative influence phenomena. Normative influence is an influence based on the desire to conform to the expectations of others (Kaplan, Miller, 1987). This statement may hold true in experimental circumstances significant to the extent of any subconscious (or even conscious) perspective the participants may have regarding the quality of the authors for each of the three articles that were under evaluation in the study. Simply put, one of the underlining issues which this thesis is aiming for would be this very normative “influence” that many individual research participants may want to “conform” to in terms of the number of authors who were in charge of the study.

The repute of the “group” in this dispute cannot be denied. Groups are often consulted in order to make decisions about political, economic, or technical problems (Vroom & Jago, 1988). This is based on the assumption that groups possess more information than a single decision-maker and, therefore, make better decisions (Clark & Stephenson, 1989; Hollenbeck et al., 1995). For instance, many judicial systems presume that groups of jurors make better arbiters of the facts than do individual jurors. Groups are often perceived to have an advantage over individuals, because groups have access to more resources than individuals.

However, groups are also known to fail to transform their superior knowledge into higher decision quality. Factors such as social loafing, the restriction of adherence to group norms and conflict in individual thought may lead to a scattered, confusing article/abstract with excessive ideas and obvious conflict in theory. It is impossible to determine however which participants have this knowledge and whether the number of

authors was actually a greater factor than the content of the abstracts. Despite this, due to the success of counterbalancing – by shuffling up the order of the three abstracts and using every possible variation of article ordering equally - it can be ruled out that the abstracts were weighed purely by its content “quality”.

The multi-level theory suggested by Hollenbeck et al. (1995) is a crucial scientific contribution that can provide a theoretical model to the still-weak body literature of hierarchical decision-making. The multi-level theory was developed to address the gap in theory and prudently explain decision-making performance for hierarchical decision-making teams with distributed expertise by identifying a small core set of variables that primarily drives accuracy in decision-making.

This theory identified three core variables (team informity, staff validity, and hierarchical sensitivity) that are central to accuracy in decision-making and mediate the effects of noncore variables on accuracy in hierarchical teams with distributed expertise. Despite the fact the participants were approached individually, we can still consider them as part of a unit or “team” as they were “officiating” as part of a voting board gearing towards the publication of articles. The multi-level theory is one of the notable perspectives in the area of team decision-making. Unfortunately, however, some aspects of the theory need refinement. In particular, team informity is not appropriate as a core variable and the operational definition of staff validity is problematic.

Despite the aforementioned limitation of “inexclusivity” of the participants, this limitation somewhat broadens the scope of things especially perhaps in a newfound

interest in implicit processing – a factor further liberated up by not questioning the participants of their awareness of the names and number of authors - during evaluation of literature yet limits the study's significance to the evaluation process of psychology literature. The present study focused on finding out the effect of the number of authors on the evaluation of abstracts for journal submission. The primary conclusion of the results was that there is a significant effect of number of authors on the ratings. It is worth mentioning there were a number of other independent variables examined, including the ordering of the articles and acceptance rates of the journal that were presented for the evaluation.

Although these variables were inconclusive, the factors of shame and guilt were also potential factors for effect. Yet although shame and guilt are usually classed together without much deliberation in the social spectrum, "...Although these variables were inconclusive, the factors of shame and guilt were also potential factors for effect. Before continuing on this sub-topic, it is essential to clearly draw the line between the two. Shame and guilt are often classed adjacent to each other in the normative social spectrum as well as textbook definitions. However, many psychologists explain otherwise – that shame is a rather self-directed, introspective emotion conditioned by cultural, moral upbringing and feels something along the lines of painful embarrassment derived from self-pride. Hypothetically, a “mistake bringing shame to all my family”; "suffering the ignominy of being sent to prison” could be analogous to shame. (Kaufman, 1996). Guilt (although equally modest in nature) on the other hand, is sometimes reputed to be less introspective and differentially outgoing in nature (Bradshaw, 1988). "I shouldn't have done that to him",

"I feel sorry to/for him". Again the sorry "to" and "for" is also discriminable, depending on the act triggering such emotions. Hereby, guilt acts as a positive reinforcer towards the independent variable in this present study. In this way and this way only may the "guilt" factor manifest itself in a participant's making of a decision to vote for a higher/multiple number of authors than a single author, or two despite the irrelevance of the authors to the actual content. Therefore, we concluded shame is not and cannot be an intrinsic factor worth mentioning considering the purpose of this study..."

We initially assumed participants may feel more guilt in deciding against a higher number of authors when evaluating an article. These factors were not measured, since we assumed post-evaluation self-report questionnaires would ultimately be ineffective in portraying the actual subconscious process during the evaluation. The fact that these variables were not measured neither adds to, nor takes away from the power of the study, since there are too many two-ended possibilities for conclusion. To prove this point – we came across a number of participants who honestly expressed that they did not think about the number of authors. Had these comments been measured, one would tend to conclude it would have a balancing effect on any further possible variable which would have subtracted from the main findings of the influence of the number of authors.

There are also other factors which should be discussed not only for the significant area of the results but not the non-significant areas. On a purely theoretical plane, the level of significance of the number of authors may not have been as high as it could have been if not for individual differences in awareness thoroughness of their reading of the abstracts. Some participants may have been busier or more pre-occupied than others

during their evaluation, which may have distorted their feedback. Again, re-testing, a more flexible system in scheduling participation coupled with a more valuable reward system would have counteracted against these limiting effect.

Also, some individuals are known to favour single or lone authors more than numerous authors. These beliefs may be founded upon their faith in the power of independent study – the type of study which may be perceived to allow for more creativity as well as restrict social loafing and restriction of individual authors' maximum potential due to the pressure to adhere to the norm of the group. It is interesting to probe into this matter, as it does make sense considering that most of the individuals who participated in this study were tertiary students who would have spent most of their educational life working alone when doing writing similar articles or essays. This idea is logically placed opposite to the main hypothesis of this study – that a higher number of authors would render positive effect upon evaluation. There is always a counter-theory to every theory and this study is no exception.

### *Scientific Implications*

The evidence we have collected has potential implications for the number of authors on evaluation of psychology abstracts. However, considering the flexibility of the rather simply designed template for evaluation scientific journals, articles as well as magazine articles and newspaper editorials could also be included for future research into the

number of authors on its evaluation and acceptance processes. For instance, there has been considerable literature on “unfair” over-sighting of political science publications and weakness of feedback (and reasoning) on rejected articles for authors (e.g. Borer, 1997). Each voting system, especially peer evaluation systems could be directly and indirectly impacted once research into this matter broadens and gains sizeable attention.

Due to the successful results of this study, the number of authors could be considered as a “distractor variable”– a new kind of term which can and should be used in future studies in the vicinity of the topics of normative norms in all kinds of literature evaluation.

### *Conclusion*

In conclusion, it is our opinion that, whilst surrounded by other issues, many people are still mostly interested in journal ratings because either on an implicit or an explicit level, the authors (the number of authors) are being evaluated as well as the journals. This relationship would be significant only up to the level of contrast the journals are not being evaluated because of this “distractor variable”. Certainly more work is necessary to unearth more of the mechanism of the relationship between the number of authors and evaluation scores as the current study broadened our understanding of the normative and unnormative influences related to the variable. Awareness of the current data also broadens the potential breadth of future research into potential factors for maturity of both the process of journal evaluation and presentation of journal submissions.

---

# References

- Abraham, A. (1973). A model for exploring intra and inter-individual processes in groups. *International Journal of Group Psychotherapy*, 23, 3-22.
- Abraham, A. (1974-1975). Processes in groups. *Bulletin de Psychologie*, 28, 746-758.
- Abrams, D., Marques, J. M., Bown, N., & Henson, M. (2000). Pro-norm and anti-norm deviance within and between groups. *Journal of Personality and Social Psychology*, 78, 906-912.
- Augustinova, M., Oberle, D., & Stasser, G. L. (2005). Differential access to information and anticipated group interaction: Impact on individual reasoning. *Journal of Personality and Social Psychology*, 88, 619-632.
- Baker, S. M., & Petty, R. E. (1994). Majority and minority influence: Source-position imbalance as a determinant of message scrutiny. *Journal of Personality and Social Psychology*, 67, 5-19.
- Baxter, G. (2009). On exclusivity and inclusivity of literary journals. *Circa*, 130.
- Borer, D. (1997). The Ugly Process of Journal Submissions: A Call For Reform. *PS: Political Science and Politics*. 30, 558-560.

- Brackbill, V. & Korton, F. Journal Reviewing Practises: Authors' and APA members' suggestions for revision. (1970), *American Psychologist*, 25, 937-940.
- Bradshaw, J. (1988). *Healing the Shame That Blinds You*. HCI.
- Burstein, E. and A. Vinokur.
- - (1975) What a person thinks upon learning he has chosen differently from others: nice evidence for the persuasive-arguments explanation of choice shifts. *J. of Exper. Social Psychology* 11: 412-426.
- - (1973) Testing two theories about group induced shifts in individual choice,. *J. of Exper. Social Psychology* 9 : 123-130.
- Chase, J. (1973). Normative criteria for scientific publication. *American Sociologist*, 8, 187-191.
- Cialdini, (2003). R. Crafting Normative Messages to Protect the Environment. *Current Directions in Psychological Science*.
- Cialdini, R., Reno, R., Kallgren, C. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, pp. 1015-1026.
- Clark, R.d., III, Crockett, W.H., and Archer, R.L. (1971). Risk-as-value hypothesis: the relationship between perception of self, others, and the risky shift. *J.Pers. soc. Psychol.* 20, 425-429.
- Clark, N. K., and Stephenson, G. M. (1989). Group remembering. In P. B. Paulus, ed., *Psychology of Group Influence*, 3rd edition. Hillsdale, NJ: Erlbaum.

- 
- Ellemers, N., van Rijswijk, W., Roefs, M., & Simons, C. (1997). Bias in intergroup perceptions: Balancing group identity with social reality. *Personality and Social Psychology Bulletin*, 23, 186-198.
  - Garvey, W.D., & Gottfredson, S.D. (1976). Changing the System: Innovations in the interactive social system of scientific communication. *Journal of information processing and management*, 1976, 12, 165-176.
  - Kang, Mincheol (2010). A proposed improvement to the multilevel theory for hierarchical decision-making teams. *Journal of management and organization*. [http://www.findarticles.com/p/articles/mi\\_qa5425/is\\_201003/ai\\_n53932103/](http://www.findarticles.com/p/articles/mi_qa5425/is_201003/ai_n53932103/)
  - Kaufman, G. (1998). *Psychology of Shame*. Churchill Livingstone.
  - Lamm, H. , Trommsdorff, G., Rost-Schaude, E. (1972). Self-image, perception of peers' risk acceptance and risky shift. *European Journal of Social Psychology*, 2, 255-272.
  - Link, A. (1998). US and Non-US Submissions: An Analysis of Reviewer Bias, *JAMA*, 246-247.
  - Perkins, H. W. (2002). Social norms and the prevention of alcohol misuse in collegiate contexts. *Journal of Studies on Alcohol*, 14, 164-172.
  - Petty, R.E., Fleming, M.A., & Fabrigar, L.R. (1999). The review process at PSPB. *Journal: personality and Social Psychology Bulletin*.
  - Rüsç, N., Corrigan, P., Bohus, G., Jacob, R., Brueck, K., Lieb, (2005),
  - Measuring shame and guilt by self-report questionnaires: a validation study. *Psychiatry Research*, 150, 313-325.

- Sherif, M. (1935). A study of some social factors in perception. *Arch. Psychol.* 187.
- Symonds, R., Gemmell, N., Braisher, K., Gorringer, L., Elgar, M., (2006). Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance. *PLoS ONE*, 127
- Tangney and Dearing (2003). *Shame and Guilt*. Guilford Publications.
- Tregenza, T. Gender bias in the refereeing process. (2002), *Trends in Ecology and Evolution*.
- Vroom, V.H. & Jago, A.G. (1988). *The new leadership: Managing participation in organizations*. Englewood Cliffs, N.J.: Prentice Hall.
- Wolff, W.M. (1970) A study of criteria for journal manuscripts. *American Psychologist*, , 25, 636-639.