

# “Face Validity” Differences between Locally Published and Highest Level International Journals Based on the In-Text Citations Approach

Siniša Subotić

[sinisasub@gmail.com](mailto:sinisasub@gmail.com)

University of Novi Sad, Faculty of Philosophy, Dpt. of Psychology, Dr Zorana Đinđića  
2, 21000 Novi Sad (Serbia)  
NGO “Persona”, 78000 Banjaluka (Bosnia and Herzegovina)

## Abstract

Locally published journals are often in a disadvantage when being compared to the highest level international journals using popular approaches such as the impact factor. This research explores an alternative approach, which focuses on the internal article’s “face validity” characteristics, based on the in-text citations measures. The sample comprised 157 empirical articles (psychology = 65, kinesiology = 32, psychiatry = 26, pedagogy = 23, and others = 11) from the journals published in Serbia in the period from 2004 to early 2010, which were compared to the matched 157 articles from the highest level international (English language) journals. ULS factor analysis conducted on the various in-text citations variables revealed two factors: *Integration* (containing variables such as: number of citations and references, number of citations that occur in multiple sections etc.), and *Superficiality* (containing: missing, imprecise and secondary citations). Large differences were obtained between the groups on both factors: Integration was significantly higher in the international group ( $d = 1.49$ ), and Superficiality in the Serbian group ( $d = 1.10$ ). Ward’s cluster analysis conducted on the two factors’ scores revealed that the articles can be classified in two clusters, with the first containing 117 international, and 9 Serbian articles (psychology = 4, psychiatry = 2, kinesiology = 2, and pedagogy = 1), and the second containing 148 Serbian, and 40 international articles (kinesiology = 16, psychiatry = 12, psychology = 5, pedagogy = 4, and others = 3), with the cluster distribution difference between the local Serbian and international articles being significant:  $\chi^2(1, n = 314) = 154.61, p < .001, \phi = -.70$ . Strong differences between the clusters on both factors were obtained, showing higher scores on Integration in cluster one ( $d = 2.14$ ), and higher scores on Superficiality in cluster two ( $d = 0.93$ ), thus the clusters were named as a *Higher tier* and a *Lower tier (of articles)*. It should also be noted that there is a certain trend of improvement in Serbian articles on the Integration di-

mension over the observed time span ( $\rho = .29, p < .001$ ). In general, the results confirm that the in-text citations based approach, and subsequently derived latent dimensions can be used to differentiate locally published and the highest level international journal articles across various scientific disciplines, and that on the average, most of the Serbian articles fall behind most of the international articles, both on the grounds of internal citation integration, and citation superficiality and mistakes. These findings directly suggest the areas of improvement for journals' editorial boards, reviewers, and authors.

**Keywords:** local journals; journal quality; citation analysis; face validity; social sciences; Serbia; factor analysis; cluster analysis; impact factor

## Introduction

Although there are some differences and exceptions between disciplines, scientific journal articles are the main and increasingly important source of scientific communication and diffusion of knowledge ([Larivière, Archambault, Gingras, & Vignola-Gagné, 2006](#)). Therefore, determining and continuously monitoring the quality of journals and articles is of the utmost importance. The most (in)famous and widely used tool for this purpose is definitely the impact factor (see [Althouse, West, Bergstrom, & Bergstrom, 2009](#); [Garfield, 2006](#)). However, the impact factor tends to vary across fields and over time ([Althouse et al., 2009](#)), making it non-optimal for interdisciplinary comparisons. Furthermore, although it can be considered as a valid measure for the quality assessment of scientific journals, it is not a valid measure for assessing the quality of individual papers and scientists ([Opthof, 1997](#)).

The usage of the impact factor is especially problematic when evaluating so called locally published journals (i.e. journals published outside of the most developed countries, and often not in English language), because they are generally underrepresented in the major citation indexing services (Archambault, Vignola-Gagné, Côté, [Larivière, & Gingras, 2006](#); [Sipka, 2004](#)), to the extent that the impact factor is arguably a completely invalid measure of these journals' quality ([Sipka, 2004](#)). Since locally published journals are a necessary complement to the highest level international journals, as they deal with the region, culture and language specific topics and serve as an educational and training tool for young researchers ([Sipka, 2004](#)), it is necessary to try and keep local journals at

least somewhat competitive with their international counterparts regarding quality. Therefore, different, less “unfair” ways of comparison besides the impact factor should be explored.

There are various possible ways to approach the evaluation of the local journals’ (formal) quality besides the impact factor (see [Sipka, 2004](#)). One of the less known concepts is in-text citations counting, which has recently been demonstrated as a successful way to improve the accuracy of assessing scientific contribution ([Hou, Li, & Niu, 2011](#)), and is a simple alternative to the complex context analysis approaches (e.g. [Huang et al., 2006](#)). Similarly to the impact factor, a possible weakness of this concept is large variability in the authors’ motivations to cite a source, as well as in the purposes that citations serve ([Bornmann & Daniel, 2008](#); see also [de Bellis, 2009](#)). However, none of this seems to call into question the role of a citation as a reliable measure of impact ([Bornmann & Daniel, 2008](#)).

[Šipka \(1996\)](#) proposed an in-text citations approach that takes into account not only the number of individual citations, but also the citations’ “re-coverage” (i.e. multiple occurrence) in different sections of the article, the citations’ age, etc. This approach, thus, focuses on the article’s internal citation characteristics, which arguably captures some of its immanent formal quality characteristics – [Šipka \(1996\)](#) refers to this as the “face validity”. [Šipka \(1996\)](#) successfully tested the “face validity” idea, showing that various in-text citations based variables group into two discriminant functions: *Article’s integrity* and *Thoroughness vs. superficiality*, and that based on them meaningful and reasonably precise discrimination between highest level international and locally published Serbian psychology articles can be achieved.

Considering the general impact factor based comparisons’ limitations ([Althouse et al., 2009](#); [Opthof, 1997](#)), and the specific limitations for local vs. international journals comparisons ([Sipka, 2004](#)), the problem of this research is to explore the alternative, in-text citations approach, and its usefulness in comparing the formal quality of local and the highest level international journals. Specifically, the goal is to revise and expand [Šipka’s \(1996\)](#) ideas, and test if such in-text citations based “face validity” measures can be used to compare the formal quality of local and interna-

tional journals from multiple scientific disciplines, since Šipka's (1996) findings were limited to the field of psychology only. If the results are satisfactory, these procedures could be considered as tools for monitoring and improving local journals' quality.

## Method

### *Sample and procedure*

The sample comprised 157 scientific articles from the journals published in Serbia (i.e. locally published journals) in the period from 2004 to early 2010, which were compared to the matched 157 articles from the highest level international (English language) journals<sup>1</sup>. Following Šipka's (1996) original procedure, only the empirical, IMRAD format journal articles (i.e. articles containing Introduction, Method, Results and Discussion sections) were included in the research. This also included several Serbian ( $n = 7$ ) and international ( $n = 3$ ) articles that had slight deviations from IMRAD (but were logically equivalent to it), or that had Results and Discussion as a joined section. The latter was more characteristic for Serbian ( $n = 50$ ) than the international articles ( $n = 3$ , with  $n = 9$  "hybrid" structure articles<sup>2</sup>). Since these aspects were impossible to control, analogous to Šipka (1996), Results and Discussion were treated as one section for all the articles (referred to as: Results & discussion).

The selection of the scientific disciplines to include journals from was dictated by the Serbian sample. Because of various differences between scientific disciplines (Althouse et al., 2009; Archambault, et al., 2006; de Bellis, 2009; Fischer, Tobi, & Ronteltap, 2011; Larivière, et al., 2006; Martin, & Sugarman, 2009), the research was deliberately limited to social sciences journals (*humanities*, i.e. literature, philosophy etc. were considered, but were dropped due to the low number of empirical articles), with the addition of cross-discipline fields, such as psychiatry and kinesiology. All of the articles from the selected time span available on the Serbian Citation Index (Šipka, 2005) in full, and lexically searchable text (i.e. not "scans") were considered, minus the journals that had less than two publications in total, and the outliers ( $n = 3$  articles had to be eliminated because they contained no in-text citations). This left 65 psychology, 32 kinesiology, 26 psychiatry, 23 pedagogy and 11 other<sup>3</sup> arti-

cles; 26 were written in English, 16 in English and Serbian, and the rest were available in Serbian only.

The international group comprised the same exact number of articles for every field, while also being controlled for a year of publication. These articles were selected as a deliberate, heterogeneous sample<sup>4</sup>; the full list of included journals is available on request.

In-text citations were counted with respect to the articles’ sections.

### *Data analysis*

All the analyses were conducted in R environment ([R Development Core Team, 2005](#)), using FACTOR ([Lorenzo-Seva, & Ferrando, 2006](#)), or manually. In-text citations counting was done manually, and reference parsings were done by the CEON Parser (see [Šipka, 2005](#)).

## **Results**

### *Variable preselection*

Initially, following and expanding upon [Šipka’s \(1996\)](#) research, a larger pool of variables was considered: number of citations (total number, and number and proportion in each of the articles’ major sections), number of references, coefficient of repetition (number of citations / number of references), number and proportion of sources cited various times (e.g. only once), number and proportion of missing citations (i.e. cited in text, but not on a reference list) and references (i.e. appearing on a reference list, but not cited in text), count of references occurring in two of the article’s sections ( $\cap$ ), number of secondary citations (i.e. “as cited in”), number of imprecise citations, various measures of citations’ age, etc.

However, a lot of variables were omitted due to no significant contribution (e.g. citations’ age measures<sup>5</sup>), low discriminativity (e.g. coefficient of repetition), multicollinearity and / or curvilinearity issues (all the proportion based measures were severely problematic, except the proportion of missing references, which behaved better than its raw number counterpart). Hence, after a systematic preselection process and trial factor analyses, only the ten most optimal variables were kept in the final model (see the *Latent dimensions* subsection).

### *Latent dimensions*

The ten selected variables were analyzed using an unweighted least squares (ULS) factor analysis. Optimal parallel analysis ([Timmerman & Lorenzo-Seva, 2011](#)) based on  $n = 1000$  permutations suggested one (95<sup>th</sup> percentile criteria) or two (mean criteria) significant factors. However, the *hull* method ([Lorenzo-Seva, Timmerman, & Kiers, 2011](#)) favoured the two-factor solution, and since it was also easier to interpret it was selected as the final one<sup>6</sup>. The factors were rotated using *promin* rotation. The results of the factor analysis are presented in [Table 1](#).

The first factor comprises variables concerning citing (and referencing) more, and citing the same sources in multiple sections. While (due to a different variable preselection) not identical to [Šipka's \(1996\)](#) first discriminative function (*Article's integrity*)<sup>7</sup>, this factor is directly analogous to it, probably representing the citation characteristics of a well grounded and integrated scientific text. Hence, the factor is named *Integration*.

On the surface, the second factor only mildly resembles [Šipka's \(1996\)](#) second discriminative function (*Thoroughness vs. superficiality*). However, they seem to share a similar gist, which is arguably better represented in this solution, due to the inclusion of more error measuring variables. Basically, this factor reflects citational mistakes, with some tendencies towards lower thoroughness (i.e. lower literature coverage); thus, the factor is named *Superficiality*.

### *Group differences*

Large differences ([Cohen, 1992](#)) were obtained between the two groups of journals on both factors' scores: Integration was significantly higher in the international journals group ( $t(239.732) = -13.16, p < .001, d = 1.49$ ), and Superficiality in the local Serbian journals group ( $t(194.046) = 9.65, p < .001, d = 1.10$ ).

To explore the “grouping tendencies” of the articles, Ward's hierarchical cluster analysis was conducted, using the two factors' scores. Multiple validation criteria (e.g. Silhouette and Connectivity; see [Handl, Knowles, & Kell, 2005](#)) suggested that the two-group solution is the most optimal. The first cluster contained 117 international, and only 9 Serbian articles

**Table 1. Factor matrix**

Variable	<i>Integration</i>	<i>Superficiality</i>	$h^2$	Variance
Number of citations	.99 (1.00)		1.00	1740.32
Number of references	.88 (.93)	(-.41)	.90	518.95
Introduction $\cap$ Results & discussion	.75 (.74)		.56	30.33
Introduction $\cap$ Method	.71 (.65)		.47	5.27
Number of single citations	.68 (.74)	(-.42)	.61	249.47
Method $\cap$ Results & discussion	.61 (.57)		.34	4.00
Missing references (proportion)	-.33 (-.35)		.13	.02
Number of missing citations		.56 (.54)	.30	1.66
Imprecise citations <sup>8</sup>		.56 (.54)	.29	.89
Number of secondary source citations		.55 (.54)	.29	6.60

Notes: Loadings in front of the brackets are factor pattern loadings, while the numbers inside the brackets are factor structure loadings (values  $<.32$  were suppressed); Variables with  $\cap$  refer to a count of citations that occur in two of the article’s sections (i.e. citations’ “re-coverage”);  $h^2$  = communality after extraction; KMO = .72; *Integration* eigenvalue = 4.20; *Superficiality* eigenvalue = 1.69; Variance before the rotation = 58.85%; Factors are negatively correlated in a lower medium intensity (Cohen, 1992):  $r = -.25, p <.001$ <sup>9</sup>

(psychology = 4, psychiatry = 2, kinesiology = 2, and pedagogy = 1), with two of them being in English, and the second cluster contained 148 Serbian, and 40 international articles (kinesiology = 16, psychiatry = 12, psychology = 5, pedagogy = 4, and others = 3). The cluster distribution difference between the local Serbian and international articles is significant and strong (Cohen, 1992):  $\chi^2(1, n = 314) = 154.61, p <.001, \phi = .70$ . Large differences (Cohen, 1992) between the clusters on both factors were obtained, with higher scores on Integration in cluster one ( $t(169.433) = 16.51, p <.001, d = 2.14$ ), and higher scores on Superficiality in cluster two ( $t(262.577) = -9.27, p <.001, d = 0.93$ ), thus the clusters were named as a *Higher tier* and a *Lower tier (of articles)*.

It should also be noted that there is a lower medium ([Cohen, 1992](#)) trend of improvement in Serbian articles on the Integration dimension over the observed time span ( $\rho = .29, p < .001$ ).

## Discussion

The main goal of this research was to explore the in-text citations based approach ([Hou et al., 2011](#); [Šipka, 1996](#)), as an alternative to the non-optimal impact factor based assessment of the locally published journals' quality ([Sipka, 2004](#)). In this regard, a “revived” and slightly modified [Šipka's \(1996\)](#) “face validity” concept turned out to be a useful tool for comparing the local (Serbian) and the highest level international scientific articles from several fields.

Two meaningful latent in-text citations dimensions (reasonably similar to the functions obtained by [Šipka, 1996](#)) concerning (1) internal citation integration and (2) citation superficiality and mistakes were obtained, and based on them it was shown that local Serbian articles fall strongly behind the international ones on both grounds, with only a handful of them reaching the higher tier of articles. While these findings should be replicated on other fields' (e.g. natural sciences & engineering) articles, and on local journals other than Serbian to allow for broader generalization, and while they can only be safely interpreted as differences in formal (citation) quality, they nevertheless suggest several areas of improvement (mostly) for local journals. First off, even though there is already a certain trend of improvement over time, that trend is not sufficiently large, thus local (Serbian) journals' editors and reviewers should explicitly insist on higher literature coverage and better integration of ideas and arguments between the sections of submitted manuscripts (i.e. aspects indirectly tentatively measured by the Integrity factor). Authors should also make a greater effort to follow those guidelines. Furthermore, more effort on everyone's part should be put into discouraging the use of secondary sources, and also the detection and correction of mistakes of any kind, including citation errors, even after publication – while this might seem trivial, the rate of error correction and retraction is actually a strong sign of journals' quality ([Fang & Casadevall, 2011](#)). Thus, journals should adopt a strict(er) peer and editorial review process, which would include citations and other formal mistakes checking procedures. It can be argued that higher formal

strictness during reviews should probably enhance the educational and training roles of local journals (Šipka, 2004), and complement other aspects of quality control, since “face validity” is necessary, but not enough to alone guarantee quality (Šipka, 1996).

Besides the already mentioned necessities for replications, since “face validity” is more of a component than a correlate of quality (Šipka, 1996), further research should aim to take some additional aspects of quality into consideration. Furthermore, some limitations of this research should be revised, e.g. more advanced mathematical procedures (that are not limited by curvilinearity, thus allowing the integration of more potentially useful variables) should be explored, and some additional indicators should also be considered. Finally, between / within disciplines specificities and “face validity” norms should be determined.

### Acknowledgements

The author wishes to thank P. Šipka for his logistic help, and O. Dakić for help with English.

### References

- Althouse, B. M., West, J. D., Bergstrom, C. T., & Bergstrom, T. (2009). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science and Technology*, 60(1), 27-34. [doi:10.1002/asi.20936](https://doi.org/10.1002/asi.20936).
- Archambault, É., Vignola-Gagné, É., Côté, G., Larivière, V., & Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics*, 68(3), 329-342. [doi:10.1007/s11192-006-0115-z](https://doi.org/10.1007/s11192-006-0115-z).
- Bornmann, L., & Daniel, H. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45-80. [doi:10.1108/00220410810844150](https://doi.org/10.1108/00220410810844150).
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. [doi:10.1037/0033-2909.112.1.155](https://doi.org/10.1037/0033-2909.112.1.155).
- de Bellis, N. (2009). *Bibliometrics and citation analysis: From the Science citation index to cybermetrics*. USA: The Scarecrow Press. Retrieved from [http://203.128.31.71/articles/0810867133\\_LIS.pdf](http://203.128.31.71/articles/0810867133_LIS.pdf)

- Fang, F. C., & Casadevall, A. (2011). Retracted Science and the Retraction Index. *Infection and Immunity*, 79(10), 3855-3859. doi:[10.1128/IAI.05661-11](https://doi.org/10.1128/IAI.05661-11).
- Fischer, A. H., Tobi, H., & Ronteltap, A. (2011). When Natural met Social: A Review of Collaboration between the Natural and Social Sciences. *Interdisciplinary Science Reviews*, 36(4), 341-358. doi:[10.1179/030801811X13160755918688](https://doi.org/10.1179/030801811X13160755918688).
- Garfield, E. (2006). The History and Meaning of the Journal Impact Factor. *Journal of the American Medical Association*, 295(1), 90-93. doi:[10.1001/jama.295.1.90](https://doi.org/10.1001/jama.295.1.90).
- Handl, J., Knowles, J., & Kell, D. B. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics*, 21(15), 3201-3212. doi:[10.1093/bioinformatics/bti517](https://doi.org/10.1093/bioinformatics/bti517).
- Hou, W., Li, M., & Niu, D. (2011). Counting citations in texts rather than reference lists to improve the accuracy of assessing scientific contribution. *BioEssays*, 33(10), 724-727. doi:[10.1002/bies.201100067](https://doi.org/10.1002/bies.201100067).
- Huang, S., Yu, Y., Xue, G. R., Zhang, B. Y., Chen, Z., & Ma, W. Y. (2006). TSSP: Multi-features based reinforcement algorithm to find related papers. *Web Intelligence and Agent Systems*, 4(3), 271-287. Retrieved from <http://iospress.metapress.com/content/84erkhh60q0vvnwfw/>
- Larivière, V., Archambault, É., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), 997-1004. doi:[10.1002/asi.20349](https://doi.org/10.1002/asi.20349).
- Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, 38(1), 88-91. pmid:[16817517](https://pubmed.ncbi.nlm.nih.gov/16817517/). doi:[10.3758/BF03192753](https://doi.org/10.3758/BF03192753).
- Lorenzo-Seva, U., Timmerman, M. E., & Kiers, A. L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, 46, 340-364. doi:[10.1080/00273171.2011.564527](https://doi.org/10.1080/00273171.2011.564527).
- Martin, J., & Sugarman, J. (2009). Does Interpretation in Psychology Differ From Interpretation in Natural Science. *Journal for the Theory of Social Behaviour*, 39(1), 19-37. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-5914.2008.00394.x/abstract>.
- Opthof, T. (1997). Sense and nonsense about the impact factor. *Cardiovascular Research*, 33(1), 1-7. doi:[10.1016/S0008-6363\(96\)00215-5](https://doi.org/10.1016/S0008-6363(96)00215-5).

- R Development Core Team (2005). R: A language and environment for statistical computing (Version 2. 14. 1) [Computer Software]. Austria, Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>.
- Šipka, P. (1996). Procena "očigledne valjanosti" naučnih radova iz oblasti psihologije na osnovu navoda u tekstu [Estimation of "Face Validity" of psychological research papers based on in-text citations]. *Psihologija*, 29(4), 417-428. Retrieved from <http://scindeks-clanci.ceon.rs/data/pdf/0048-5705/1996/0048-57059604417S.pdf>
- Šipka, P. (2004). Vrednovanje časopisa na osnovu bibliometrijskih pokazatelja: Opis modela i implikacije za uređivački rad [The evaluation of journals based on bibliometric indicators: Description of a model and implications for editorial work]. *Srpski arhiv za celokupno lekarstvo*, 38-49 (special issue). Retrieved from [http://ceon.rs/pdf/vrednovanje\\_casopisa\\_na\\_osnovu\\_bibliometrijskih\\_pokazatelja.pdf](http://ceon.rs/pdf/vrednovanje_casopisa_na_osnovu_bibliometrijskih_pokazatelja.pdf)
- Šipka, P. (2005). The Serbian Citation Index: Context and content. In: *Proceedings of ISSI 2005 - 10th International Conference of the Society for Scientometrics and Informetrics*, Stockholm, SE July 24-28, 2005 (pp. 710-711). Retrieved from [http://ceon.rs/pdf/Sipka\\_SCIndeks\\_proceedings.pdf](http://ceon.rs/pdf/Sipka_SCIndeks_proceedings.pdf).
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209-220. doi:10.1037/a0023353.

## Notes

- [1] Šipka (1996) tried to differentiate between the “academic” and “non-academic” local journals, but his results showed some problems with such a distinction. Therefore, it was not considered here.
- [2] I.e. containing several Studies, but some with Results and Discussion separated, and some with them joined.
- [3] Containing mostly articles on geography and economy.
- [4] A random selection was not optimal, because some sort of rudimentary content control was necessary. For example, while pedagogy exists in Serbia, it is not an internationally established discipline, so the international subsample had to be selected out of the educational psychology articles dealing with specific topics, such as educational

measurement and tests, etc. Only titles, abstracts and key words were used in the content matching process. Furthermore, the Serbian group usually had one dominant field-specific journal, dealing with a wide range of topics, while international journals are generally more specialized – to match this, international articles were roughly proportionally (quasi-randomly) selected from multiple (two to three) highly ranked journals from each field. Following the described methodology, roughly 50% more international articles were preselected for each field (controlling for a year of publication); from this preselected pool, the final articles were selected randomly.

- [5] Even though irrelevant here, the citations' age should probably be standardly considered in future comparisons.
- [6] Factor invariance was impossible to check due to sample size, but it is plausible to assume that, while relevant on the Serbian and joined sample, the second factor would be suppressed on the international sample alone.
- [7] Note that this function also had a proportion of missing references negatively loading on it.
- [8] This variable does not include spelling errors, but rather cases where a cited source has multiple authors, but only one name is cited in-text, or when there are multiple works by the same author(s), from the same year, but they are not differentiated in text (i.e. "a, b, c..." were not used), etc. Note that this variable was truncated to 5 points (0, 1, 2, 3, 4+ imprecise citations) for two reasons: 1) 96.5% of the articles had a score of 3 or less, with a handful of articles with the extreme values; 2) in certain cases, the exact number of imprecise citations could not be precisely determined (this was always a case with articles already having a definite score of at least 3). Thus, coding the mentioned "problematic" articles simply as "4" was a convenient way to deal with those issues.
- [9] Note that this correlation is not a result of a proportion of missing references loading on the first factor ( when that variable is omitted, the correlation drops off only for .005).