Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles

Sinisa Subotic
NGO 'Persona', Bosnia and Herzegovina and 'CEON/CEES', Serbia

Bhaskar Mukherjee
Department of Library and Information Science, Guru Ghasidas University, India

Abstract
The aim of this study was to conduct a unified investigation of various, previously mostly individually studied scientific article title characteristics, like: title length, type, amusement and pleasantness, and specific title 'markers' (e.g. colons, attention-grabbing words etc.) in relation to subsequent article citation and download rates. Based on a sample of 129 psychology ScienceDirect's Top 25 Hottest Articles (i.e. highly downloaded articles) and 129 articles not appearing on the Top 25 list (i.e. less downloaded articles), we determined that the most relevant title characteristics were the title length and the title amusement/humour. The partial least squares model revealed that shorter titles were associated with more citations, but the effect was fully mediated by the journal impact, suggesting that the observed citational benefits of the shorter titles might be an artefact of some higher journal impact related attribute (perhaps editorial or peer review process). Title amusement level was slightly correlated with downloads, but with no association with citations. Additionally, downloads correlated positively with citations, and more amusing titles tended to be shorter. While these findings are limited to the psychology discipline only, our results suggest that the integrative structural approach is promising and that more research following this paradigm is needed before empirically grounded recommendations for good title writing can be given.

Keywords
Amusement/humour; citation rates; download rates; journal impact; partial least squares (PLS); psychology; ScienceDirect's Top 25 Hottest Articles; title

1. Introduction
The title is arguably the most important element of a scientific article [1–4]. It informs the readers about the content of the article and encourages them to read it [5]; it is also used for database indexing and referencing purposes [6]. It has been suggested that about 70–90% of scientific citations are likely to be copied (without sufficient reading or reading at all) from the lists of references used in other papers [7]. Commenting on this, Moore [8] states:

In that context, a title is everything: a title that contains good keywords, and expresses the finding explicitly […] is likely to be picked up and re-cited often.

However, only a few of the title-related characteristics have been studied in detail. For example, title length correlates with the number of authors [3] and with the length of the article [1, 4], with moderate correlations for most nature science journals, but limited or negative correlations for social sciences and humanities journals in both cases [3, 4].

Research regarding the association of the title characteristics with the citation rates is also limited, but it seems to be increasing in frequency. Sage and Yechiam [2] studied the attention-grabbing’ quality of the titles of two high-level...
psychology journals and its relation to the number of citations. While this phenomenon has been studied before, especially in the marketing discipline, with the data suggesting no impact on the citation count [9], Sagi and Yechiam [2] explored the attention-grabbing in its extreme form — the use of humorous/amusing titles (in psychology articles). They took into account the levels of the titles’ amusement and pleasantness, obtained via rating of those characteristics by independent judges. Their results showed that the pleasantness rating was weakly associated with the number of citations, while an extremely high level of titles’ amusement was related to fewer citations. This negative correlation cannot be attributed to the difference in the title length, pleasantness, number of authors, year of publication or article type (defined as regular article vs comment).

The title length is among more studied title characteristics in relation to citation rates. Stremersch et al. [9] found no correlation between the title length and the number of citations (in marketing journals). Jacques and Sebire [10] explored the title characteristics of the 25 most cited articles and the 25 least cited articles from three medical journals and, contrary to Stremersch et al. [9], found that the length of the title is positively correlated with the number of citations. Jacques and Sebire [10] tentatively explained this via the specificities of (now predominant) electronic searches, because:

> a longer, more comprehensive title is both more likely to contain any given search term and, therefore, be identified, and also if the title provides a clear description of the study or its findings, is also more likely to be identified as relevant on the initial search screening process.

Other researchers, such as Habibzadeh and Yadollahie [11] have also confirmed that longer titles (of medical articles) are associated with a higher number of citations, with the effect being stronger for journals with high impact factors. It should be noted that some of the findings and conclusions reported by Jacques and Sebire [10] were later questioned by Moore [8], as he suggested the possibility of the findings being highly field and discipline specific; specifically, he suggested that the finding regarding title length and its positive association with the citation rates in some cases is negative. Indeed, such findings have been obtained, for example, in biomedical journals, where articles with short titles, describing the results, are cited more often [12].

Jamali and Nikazad [1] also obtained the results partially in line with Moore’s [8] reasoning and contradicting the findings of Jacques and Sebire [10] and Habibzadeh and Yadollahie [11]. Specifically, they showed that (for medical and life sciences) articles with longer titles tended to be downloaded slightly less often than the articles with shorter titles, and that titles with a colon tended to be longer and their articles received fewer downloads and citations (but with no direct relation between length and citations). Furthermore, Jamali and Nikazad [1] found that the articles with question (interrogative) titles tended to be download more but cited less that declarative titles (i.e. titles that also state the research’s conclusion) and descriptive/neutral titles (i.e. the titles that only describe the subject of the article without revealing the findings/conclusions). Surprisingly, Ball [13] showed that the number of articles with question marks in titles has increased significantly (from 50% to over 200%) over the last four decades in physics, life sciences and medicine. Jamali and Nikazad [1] also noted that declarative titles tended to be longer than descriptive titles and downloaded and cited slightly less. The findings regarding declarative vs descriptive titles seem to contradict the recommendation for making the findings explicit in the title [8, 12, 14], as well as the already cited explanation put forward by Jacques and Sebire [10].

There are also indices of citation rates being arguably influenced by the presence of specific markers in the titles; for example, Jacques and Sebire [10] found that the presence of a colon and the presence of an acronym in the title are positively correlated with the number of citations, while the reference to a specific country in the title is negatively correlated with the number of citations. Hartley [15], however, determined that, while students and academics generally prefer titles with a colon, its presence does not affect the number of citations received (note that these findings were partially limited to the psychology discipline). Additionally, Lewison and Hartley [16] determined that both the title length and the use of colons increased over time (for science and oncology), while Buter and van Raan [17] reported that there has been no general increase of non-alphanumeric characters usage over the last 10 years, but that characters like colons or hyphens are quite common in titles and that often not including them is associated with lower (field-normalized) impact, but that there is a huge field and discipline variability, with even absent or negative correlations.

While it is clear from the cited research that there are some relationships between various title characteristics and citation rates, the question remains how and to what extent this can be utilized by researchers looking to increase the visibility of their publications. On that note, Moore [8] suggested that:

> title/citation analysis cannot per se be developed into a strategy to increase citations: those are overwhelmingly driven by other things. However, it can be useful to optimize a title in order to help an article gain its deserved reader attention.

Journal of Information Science, 40(1) 2014, pp. 115–124 © The Author(s), DOI: 10.1177/0165551513511393
Still, as previous research mostly dealt with title characteristics related to citations in an isolated fashion, it is difficult to give direct, empirically grounded recommendations on how to write a good title. Based on that limited knowledge, we can state that the following title attributes might be relevant to subsequent article citations (with subject field/discipline specificities and even temporal changes as possible moderators):

- the title length [1, 8, 10–12];
- title type, e.g. declarative, descriptive and question titles as defined by Jamali and Nikazad [1], with the possible addition of hybrid/compound titles;
- the presence of specific ‘markers’, like a colon [1, 10], an acronym or a specific country reference [10], and extreme attention grabbers, such as humour/amusement [2].

For a summary of the research findings on title characteristics related to article citations, see Table 1.

The aim of our study was to explore a unified approach of simultaneous consideration of these various title characteristics in relation to not only citation rates, but also usage data (i.e. download rates or other available download proxies) as well. We suspect that the download rates might be of a particular importance here because, on the one hand, the title characteristics seem to be related to both the download rates [1] and the citation rates of an article [1, 2, 10, 11], with indices...
that the first association is stronger [1]. On the other hand, the number of times an article is downloaded also correlates with the number of citations it receives [1, 18–20]. This suggests that there actually might be both a direct and indirect relationship (e.g., through the article’s download rates) of the title characteristics with the citation rates. In this regard, we will pay special attention to the extreme attention-grabbing property, that is, humour, as it is reasonable to expect its presence amongst the most downloaded articles’ titles, even though the relation between this phenomenon and the number of downloads has never been studied before and only data showing its weak negative association (at the extreme values) with citation rates exists [2]. We suspect that it is possible that humorous titles are downloaded more, but not necessarily cited more, or are even cited less [2]. Also, while there is no consensus regarding the direction (or even the existence) of the relationship between the title length, citations and downloads, there are indices for various possible paths by which the title length might influence downloads and citations, with multiple variables possibly playing a mediating or moderating role, like title type [1] or the journal’s impact factor [11].

While some have already adopted this ‘title + downloads + citations’ approach [1], no previous research has taken into account all of the identified relevant title characteristics and explored various possible direct and indirect paths by which the title characteristics might influence either downloads and citations or both.

2. Method

2.1. Data gathering

The number of article downloads or views has usually been considered ‘a business secret’ by most subscription-based publishers, thus this information is generally not readily available for public research purposes. While previous research dealt with this issue by taking into account download data for a small number of open access [1] or even individual (subscription) journals [18–20], we wanted to take a different approach that would allow us to study multiple subscription-based journals. Even though the precise number of downloads may not often be available, there are publicly available ‘proxies’ to downloads, such as the ScienceDirect’s Top 25 Hottest Articles, which comprises a quarterly updated list of the 25 most downloaded articles for a given journal or the subject area. This list can be viewed as a rank-ordered downloads variable that is discriminative at the high values portion of the variable, but with truncated other (i.e., middle and low) values – those values can, however, be added by including the corresponding articles that have never appeared on the Top 25 list as the 26th rank value. This would result in a downloads approximation variable with 25 ranks representing high downloads and one rank representing everything else – conceptually, this could also be viewed as the highest values–other values dichotomy. While far from perfect (especially with regards to non-existing discriminativity on non-high values), this variable can still potentially be used as a gateway to otherwise unavailable subscription journals’ article download rates.

Following the described rationale, we collected the download ranks from the Top 25 list for psychology subject area articles (joined list of all journals) for all the quartals of 2008, 2009, and 2010. We selected the psychology discipline mainly because the previous research dealing with the title amusement has been conducted on psychology articles [2]. Some articles have appeared on the list for multiple quartals and years, and some have been published in years prior to that of their list appearance; since this made it difficult to assess the appropriate citation window and also arguably increased the risk of the Matthew effect [22], we removed multiple appearing articles and articles not published in the same year as their list appearance. This left the total of \( n = 129 \) single appearing articles (from \( n = 41 \) journals) for all the quartals of the selected three years. After this, we randomly selected the same number of articles from the same corresponding journal volumes (and issues when possible), that have never appeared on the list, for a total of \( n = 258 \) articles (from 40 journals). Additional info such as the number of authors and the article length were also collected.

For the selected articles, we obtained the citation numbers from the Scopus for the next two years after their appearance on the Top 25 list (e.g., for the articles appearing on the list in 2008, we obtained the citations received in 2009 and 2010). This was done to allow for enough time to pass for an article to have a chance to accumulate citations [1] and to give the articles from different years the same citation window length. We also obtained the journal impacts and 5-year impacts (where available) from the journals’ web pages.

2.2. The title ratings and the title-related data

Seven trained judges proficient in English and science communication (as well as in the psychology discipline) rated the humour/amusement and pleasantness ratings following the original procedure and instructions proposed by Sagi and Yechiam [2]. The judges were rating these attributes on seven-point Likert scales, following the Oxford dictionary definitions, according to which Amuse means – ‘cause (someone) to laugh or smile’ and Pleasant means – ‘giving a sense
of happy satisfaction or enjoyment; friendly and likeable’. The titles were presented in the spreadsheet file in randomized order.

The title type classifications [1], presence of colons and other non-alphanumerical characters [1, 11, 17], as well as acronyms and geographical references [11], were also assessed by the principal investigators.\(^5\) We also coded for attention-grabbing words [9], extending the list to 20 indicators (e.g. new, improved, better, advanced, superior). The title length was obtained via three variables: the number of total characters, the number of characters without spaces and the number of words.

### 2.3. Statistical analyses

Because of the non-normality of most of the variables, model testing was conducted using partial least squares (PLS) analysis with the SmartPLS Version 2.0 (M3) software [23]. PLS is a method for the estimation of path models with latent constructs based on indirectly measured manifest indicators, but unlike conventional (covariance-based) modelling, PLS is a soft modelling procedure, which means that it has no strong data and residual distribution assumptions (i.e. it does not assume nor require normal distribution), it requires fewer indicators per latent construct, and can even converge on smaller sample sizes than covariance methods [24] (statistical power, however, still has to be considered). Furthermore, since it does not have covariance reproduced matrix, PLS does not rely on standard ‘indicators of model fit’ as covariance-based modelling does. This, however, comes as a trade-off, because in PLS the concept of causation must be abandoned and replaced by the concept of predictability. While this might pose a limitation in certain situations, overall flexibility of the assumptions makes PLS an excellent approach for analysing scientometric data.

### 3. Results

#### 3.1. Main model

Table 2 shows the composite reliabilities (CRs), average variances extracted (AVE) and correlations among the latent variables that were kept in the final model. These values are used to judge convergent and discriminative validity, and variable retention decisions were based on them and on the strength of the manifest relationships with either downloads or citations.

Proposed thresholds for the acceptable convergent validity for the items are CR \(\geq 0.70\) and AVE \(\geq 0.50\) [25]. While all the CR values were optimal, initially there was an issue with the low Title Amusement AVE value (0.43), which was fixed by removing two amusement ratings with outer loadings less than 0.60,\(^6\) thus reducing the total number of ratings to five. After this, all of the values indicated acceptable convergent validity. The numbers in the matrix diagonals are the square roots of the AVEs, and they are greater in all cases than the off-diagonal numbers (correlations) in their corresponding row and column, suggesting the discriminant validity of the variables. Note that Downloads is a single-item construct, which explains ‘perfect’ AVE and CR values. Finally, no high cross loadings were observed, which also confirms the convergent validity of the variables.

Pleasantness rating correlated highly with the amusement rating (\(\rho = 0.71, \ p < 0.001\)),\(^7\) while the correlation with citations was weak and only close to significance (\(\rho = 0.12, \ p = 0.060\)), with no association with downloads. Amusement and pleasantness ratings showed a significant overlap, thus we tested the models with either one or the other. Since amusement showed better associations with other variables and the higher AVE (after correction) and since the initial internal consistency for the amusement was higher than for the pleasantness (CR = 0.84 vs 0.79), only the amusement rating was kept in the final model. No other measured variable showed acceptable evaluation metrics and the strength of the

<table>
<thead>
<tr>
<th>Latent variables</th>
<th>AVE</th>
<th>CR</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Title amusement</td>
<td>0.51</td>
<td>0.84</td>
<td>0.71</td>
</tr>
<tr>
<td>Citations</td>
<td>0.93</td>
<td>0.96</td>
<td>0.08</td>
</tr>
<tr>
<td>Downloads</td>
<td>1.0</td>
<td>1.0</td>
<td>0.20</td>
</tr>
<tr>
<td>Impact factor</td>
<td>&gt;0.99</td>
<td>&gt;0.99</td>
<td>0.29</td>
</tr>
<tr>
<td>Title length</td>
<td>0.97</td>
<td>0.99</td>
<td>−0.42</td>
</tr>
</tbody>
</table>

---

Subotic and Mukherjee, Journal of Information Science, 40(1) 2014, pp. 115–124 © The Author(s), DOI: 10.1177/0165551513511393

---

\(^5\) We also coded for attention-grabbing words [9], extending the list to 20 indicators (e.g. new, improved, better, advanced, superior).

\(^6\) Fixed by removing two amusement ratings with outer loadings less than 0.60.

\(^7\) While the correlation with citations was weak and only close to significance (\(\rho = 0.12, \ p = 0.060\)), with no association with downloads.
association to be included in the final model, including the title type, which was expected to have a significant role, based on previous research [1].

The final model is shown in Figure 1. As it could be expected from the intrinsic nature of these variables, indicators for Title Length, Journal Impact and Citations are highly correlated (e.g. the number of citations received in the first year correlates highly with the citations received in the second year), which means that each of the latent variables could have been specified as a single-item construct using any of its indicators. Since, in practice, this yielded almost identical results, we have kept all of the indicators in the model.

Only the significant paths were retained in the model (i.e. Title Amusement–Citations and Title Length–Downloads showed no significant relationship, even when those relations were individually examined with no other paths present; note that these paths had enough observed statistical power to detect potential effects).

Downloads correlated positively with Citations and the effect was apparent (and of a lower medium size) even when we considered Downloads as a binary variable (0 = non-Top 25 articles, i.e. medium and low download rates; 1 = Top 25 articles, i.e. high download rates; $U = 5818$, $p < 0.001$, $r = 0.26$); also, there was still a small difference in amusement rating ($U = 6978$, $p = 0.024$, $r = 0.14$).

As suspected, Title Amusement was indeed a positive Downloads predictor, that is, the articles with more humorous titles were in fact downloaded more, but there was no correlation with the number of citations received. No relationship between the title amusement rating and citations emerged even when only the extreme (low and high) amusement ratings were compared ($U = 310$, $p = 0.97$, $r = 0.006$), which is in contrast to findings of Sage and Yechiam [2]. However, more amusing titles did tend to be shorter, and articles with shorter titles did in fact receive more citations, which is in line with some of the previous findings and observations [8, 12]. However, it appears that this relationship between the Title Length and Citations is mediated by Journal Impact, as we have implicitly expected based on previous research [11]. It is also mediated fully, that is, the Title Length–Citations path becomes completely non-significant after Journal Impact is entered into the model, while the indirect Journal Impact–Citations path is significant, which satisfies the criteria for full mediation [26].

Figure 1. PLS structural and measurement model: the standardized regression coefficient between Title Length and Citations after adding Journal Impact is in parentheses; squared multiple correlation coefficient after adding Journal Impact to the model is in parentheses; *** $p < 0.001$, $p$-values estimated using 5000 bootstrap samples. ‘Am’ variables represent the amusement ratings made by the individual judges/raters. IF = impact factor.

Note that we recoded Downloads, so that it goes from low to high values. IF, Impact factor.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$R^2$</th>
<th>Coefficient</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downloads</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citations Year 1</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citations Year 2</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title Amusement</td>
<td></td>
<td>0.72***</td>
<td></td>
</tr>
<tr>
<td>Title Length</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td></td>
<td>0.97***</td>
<td></td>
</tr>
<tr>
<td>Number of characters (without spaces)</td>
<td></td>
<td>0.99***</td>
<td></td>
</tr>
<tr>
<td>Number of characters</td>
<td></td>
<td>&gt;0.99***</td>
<td></td>
</tr>
<tr>
<td>Journal IF</td>
<td></td>
<td>&gt;0.99***</td>
<td></td>
</tr>
<tr>
<td>5-Year IF</td>
<td></td>
<td>&gt;0.99***</td>
<td></td>
</tr>
</tbody>
</table>

Downloaded from jis.sagepub.com by guest on July 14, 2015
Following Cohen’s [27] guidelines, $f^2$ effect sizes for $R^2$ statistics were calculated, using the formula: $f^2 = R^2/(1 - R^2)$. Cutoffs for small, medium and large effect sizes are 0.02, 0.15, and 0.35. The variances explained in Downloads ($f^2 = 0.04$) and Citations ($f^2 = 0.11$) before entering Journal Impact were small. Medium amount of variance was explained in Title Length ($f^2 = 0.20$) and Citations ($f^2 = 0.23$) after entering Journal Impact, while a large amount of variance was explained in the Journal Impact itself ($f^2 = 0.35$).

Additionally, we established that there is a small positive relationship between the Title Amusement ($\beta = 0.28, p < 0.001$) and Journal Impact ($R^2 = 0.08, f^2 = 0.09$), but it is fully mediated [26] by the Title Length variable. We also explored potential moderation of every single relation, including the variables not retained in the final model as moderators – no single moderation appeared to be significant.

3.2. Additional analyses

The mean title length of the examined psychology articles was mean ($M$) = 78.2 (standard deviation, SD = 32.7) characters ($M$ = 68.8 (SD = 28.7) without spaces), or $M = 10.5$ (SD = 4.26) words, with the longest title having 182 characters (158 without spaces) or 25 words, and the shortest title having 24 characters (21 without spaces) or three words.

The titles with colons (41.1%) tended to be longer to a larger extent ($t(256) = -6.31, p < 0.001, d = 0.80$) than titles without them, but there was no correlation with either citations or downloads. Longer titles also had a moderate tendency to come from longer articles ($\rho = 0.36, p < 0.001$) and a small tendency to come from more authors per article ($\rho = 0.27, p < 0.001$), which is all in line with previous research [1, 3, 4]. Furthermore, there was a small length difference between title types as described by Jamali and Nikazad [1] ($F(3, 254) = 4.28, p = 0.006, \eta^2 = 0.05$), with contrasts suggesting that the compound titles were significantly longer than the question titles ($\rho = 0.049$) and almost significantly longer than the descriptive titles ($\rho = 0.054$); however, title type showed no relation with downloads or citations. It should be noted that there was a large distribution discrepancy between the title types ($\chi^2(3, N = 258) = 398, p < 0.001, \chi = 1.21$), such as that the descriptive titles were by far the most prevalent (78.7%), followed by compound (9.7%), declarative (7.8%), and question titles (3.9%). Attention-grabbing words were slightly more prevalent in titles with colons ($\chi^2(1, N = 258) = 5.01, p = 0.02, w = 0.14$), and titles with the attention grabbers (9.3%) were judged as slightly more amusing ($U = 1811, p = 0.008, r = 0.18$) when compared with titles with no attention-grabbing words, but with no association with downloads or citations. Finally, the titles with acronyms ($U = 1033, p = 0.001, r = 0.22$) and a slight tendency of being judged as less pleasant ($U = 1444, p = 0.41, r = 0.13$) with no relation with downloads or citations.

3.3. Examples of the most and the least amusing psychology titles

Examples of highly amusing titles include ‘How extraverted is honey.bunny77@hotmail.de? Inferring personality from e-mail addresses’, ‘God: Do I have your attention?’, ‘The battle for Broca’s region’ and ‘Taking a new look at looking at nothing’. Examples of highly unamusing titles include ‘Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes’, ‘The roles of financial asset market failure denial and the economic crisis: Reflections on accounting and financial theories and practices’, ‘Tobacco use among those with serious psychological distress: Results from the national survey of drug use and health, 2002’ and ‘Measurement invariance: Review of practice and implications’.

4. Discussion

It is probably not reasonable to assume that simply changing the title of an article will increase its citations [10] and in that regard what is said probably matters more than how it is said [9]. However, optimizing the title is still an important task [8], a task that has been made difficult by the field and discipline specificity and the scarcity of the findings as well as the simplicity of the research. This study focused primarily on the latter issue, as our main goal was to try to unify most of the previously identified relevant title related variables into a single predictive model, while simultaneously exploring the idea that title characteristics might be related not only to the citations that article an gets, but also to its usage statistics, that is, downloads [1]. Several key findings emerged from this approach.

First of all, our findings support the idea of shorter titles being more effective, that is, their articles receiving more citations (specifically, the title length associated with an above median number of citations had roughly 10 ± 4 words). As previously suggested, the relationship between the title length and the number of citations appears to be highly contradictory and field- and discipline-specific. One possible reason for this incongruence in findings might be the journal impact [11], as this variable is usually not directly controlled for when examining the title length–citations relationship. Our findings fully support the notion that journal impact plays the key role here, as it seems that shorter title articles tend
to get more citations since they are published more often in higher impact journals and articles from higher impact journals tend to receive more citations. While the causality of this relationship cannot be determined yet, it is plausible to hypothesize that higher impact journals, at least in the psychology discipline, insist on or promote shorter titles (e.g. during peer review), that is, shorter titles might be an outcome of journal review policy and process. However, regardless of what the actual mechanism is and possible disciplinary specificity on the side, the direct implication of this finding for future research is that journal impact should probably always be taken into consideration when examining the title length–number of citations relationship.

When it comes to the title humour, contrary to weak negative association with citations reported by Sage and Yechiam [2], we found that the articles were not penalized nor rewarded for their title amusement level in terms of citations received, but they were downloaded slightly more often (and downloads on their own were associated with more citations). One way to look at this is that some amount of title amusement might help in a small amount with attracting a bigger audience, but then other factors play a larger role in actually (not) getting citations. Sage and Yechiam [2] suggested the possibility that humorous titles communicate a non-serious subject matter and harm credibility, but their observation of negative correlation between title humour and its article citations was based on the extreme values (i.e. low and high humour), which our findings did not replicate. This discrepancy in the two sets of findings cannot be explained at this point and requires further research, but it could be speculated that title humour acts as a trade-off, where a good article with an amusing title might do better than a good article with an unamusing title, while a bad article with an amusing title might actually do worse than a bad article with an unamusing title. Another possibility is that humorous titles simply communicate better with the general public readers, making the topics appear to have more ‘real world’ relevance. Similarly, it might spark a curiosity of students or researchers not otherwise interested in that specific topic. This might increase downloads, but it will not necessarily result in any more article citations.

Positive correlation between the number of citations and downloads suggests that the download rates ‘proxy’ that we have used is indeed an acceptable operationalization of this variable, despite already described limitations, and this finding is in line with previous research [1, 18–20]. However, it is still just a crude proxy and it could be argued that the correlation was still contaminated with the Matthew effect [22], as it is plausible that the Top 25 articles received more citations partially because of the fact that they have appeared on the list. At this point, this cannot be confirmed nor denied, and a future effort should be made to try to obtain the more precise usage data from the publishers. The feasibility of such efforts will largely depend on the willingness of publishers to grant access to such data for research purposes. Also, more precise measurement of the article downloads variable could arguably influence the strength of the correlation with the title amusement. Furthermore, the whole downloads–citations relationship is still largely unexplained and it is yet to be determined what are the factors potentially mediating or moderating it.

Surprisingly, several variables showed no significant relationship with either citations of downloads. When it comes to specific title markers like acronyms or geographical references, it is quite certain that a significantly larger sample size is required to reliably test those variables, owing to their low prevalence in the sample. The same conclusion can be offered for title types, as it turned out that descriptive titles were the most prevalent ones, thus no conclusion regarding title type differences can be drawn at this moment (and subsequently no conclusion can be given for the question mark usage as well, as this variable is ‘nested’ into the title types). However, the mere fact that descriptive titles were the most frequent ones still offers some insight into the general practice of naming psychology articles. When it comes to colons, however, there is enough statistical power to support the conclusion that it probably does not matter much if they are used or not, and this is in line with findings of Hartley [15], also partially obtained on psychology titles.

It is also important to emphasize that the included title characteristic variables seem to be related to either citations or downloads, but not both. However, Jamali and Nikazad [1] reported a weak negative correlation between the title length and downloads, while we observed no such relationship. It might be that the effect is so small that we did not detect it, or that it was lost because of the nature of our downloads proxy variable. However, it might also be that there are some other factors influencing this relationship that were not examined, for example, the extent to which the title is optimized for electronic search engines (e.g. the ratio of relevant search words included in the title vs ‘filler’ words). It would probably be wise to include such a measure in future research on this topic. Furthermore, while the title amusement was not related to citations per se, more amusing titles did tend to be shorter, and amusement was in fact in a small positive relationship with journal impact, mediated through title length. On the other hand, the title length effect on citations was mediated through journal impact. This implies that there might be a very complex mediation net between those variables that we are now only beginning to understand. Taking all this into consideration, more likely than not, the current model is largely incomplete and important pieces of information are still missing to fully explain the dynamic between the variables. These pieces might either be other title related characteristics, other scientometric variables or general variables. As per the goal of our research, we focused mostly on the title related characteristics, not on the maximal citations or downloads predictions specifically; however, the future research can go either way — focusing on the title itself,
expanding this model and integrating it into a more general citations or downloads prediction model later on, or trying to build a general citations and downloads prediction model which would include title characteristics (probably as a minor part) from the start. What does seem certain is that the optimal way to complete the model is the structural approach that we are advocating, with the focus on the simultaneous examination and integration, rather than examining the variables in an isolated fashion.

The main implication of our research for authors looking to write a good (psychology) article title would be to actually not read too much into any of the titles-related findings just yet, as we are only beginning to understand how relevant title characteristics relate to each other in an integrative context. With that being said, a tentative recommendation supported by our findings would be to preferably keep the title short and amusing, within common sense and good taste, and that colon usage probably does not matter; at this time, there is not enough grounds to give recommendations on title type and other types of specific title markers. Note that, even if it appears likely that the citational benefits of the shorter titles are an artefact of some higher journal impact-related attribute, it might still be a good idea to follow a general trend proposed by the best journals, that is, keep the title short. Keep in mind, however, that our study was a small-scale preliminary study, thus findings should be replicated on a larger sample. Also, findings and implications are strictly limited to psychology discipline11 and subscription-based articles and separate examination should be done for other scientific fields and disciplines and open access articles before more specific and generalized recommendations can be given.

Acknowledgements

The authors would like to thank all of the judges/raters for their invaluable contribution to this research.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Notes

1. Attention grabbers can be defined as words that have a special appeal because they raise attention [9].
2. Note, however, that this relationship might be inflated for open access articles, as there is a controlled trial showing that open access status tends to significantly increase article downloads, but with no citation advantages [21].
3. That is, ‘the rich get richer and the poor get poorer’, where articles that are initially highly cited in their first few years tend to maintain their higher citation rates later on. It is possible that this ‘high visibility’ owing to multiple appearances on the Top 25 list has increased the Matthew effect. In fact, it is plausible that every single Top 25 article is influenced by it to an extent; however, at this point it is almost impossible to determine this precisely.
4. For an unknown reason the January to March 2009 quartal had only $n = 12$ entries on the list.
5. Note that the question marks are not mentioned here because this variable is already included in the title type classification.
6. For confirmatory purposes, outer loadings should be $\geq 0.70$, while for exploratory purposes a $\geq 0.60$ cutoff is sufficient.
7. While these results are in line with the results of Sagi and Yechiam [2], the pleasantness–amusement correlation in this research is much stronger (i.e. high vs medium).
8. Additionally, there was also no correlation between the title humour ratings and citations even when Top 25 and non-Top 25 sub-samples were treated separately, even at the extreme values (i.e. low and high humour).
9. Thus this direct relationship is not present in the final model, since Journal Impact is entered as a mediator for the Title Length–Citations relation, after which Title Amusement–Journal Impact already loses its significance ($\beta = 0.10, p > 0.05$).
10. Note that only 2.71% (i.e. seven) articles had any kind of geographical terms in the title, and because of such a low number this variable was not explored any further. The same goes for various other specific non-alphanumerical markers, which were simply not frequent enough to be included in the analyses (e.g. hyphens in a non-compound word usage were observed in only 1.55% of cases, i.e. five).
11. As one of the reviewers pointed out, it is also quite possible that even significant within-discipline differences might exist, as psychology in particular deals with very diverse topics and specific journal differences might exist as well. In fact, Whissell [28] showed that (in psychology) there are even significant within-journal title changes over time. These are all the questions that we did not cover in this research, but they should be kept in mind when considering the implications of our findings.

References