

# Improving information quality of Wikipedia articles with cooperative principle

*Miloš Fidler and Dejan Lavbič*

Miloš Fidler and Dejan Lavbič. 2018. **Improving information quality of Wikipedia articles with cooperative principle**, Online Information Review (OIR), 41(6), pp. 797 - 811.

## Abstract

**Purpose** – The purpose of this paper is to investigate the impact of cooperative principle on the information quality (IQ) by making objects more relevant for consumer needs, in particular case Wikipedia articles for students.

**Design/methodology/approach** – The authors performed a quantitative study with participants being invited to complete an online survey. Each rater evaluated three selected and re-written articles from Wikipedia by four IQ dimensions (accuracy, completeness, objectivity, and representation). Grice's maxims and submaxims were used to re-write articles and make them more relevant for student cognitive needs. The results were analyzed with statistical methods of mean, standard deviation, Cronbach's  $\alpha$ , and ICC (two-way random model of single measure).

**Findings** – The study demonstrates that Wikipedia articles can be made more relevant for student needs by using cooperative principle with increase in IQ and also achieving higher consistency of students' scores as recent research. In particular, students in the research perceived the abstract, constructed with cooperative principle, more objective and complete as reported in recent research.

**Practical implications** – The work can benefit encyclopedia editors to improve IQ of existing articles as well as consumers that would obtain more relevant information in less reading time.

**Originality/value** – This is one of the first attempts to empirically investigate the application of cooperate principle to make objects more relevant for consumer needs and impact of this on IQ. IQ improvement evidence is provided and impacts on IQ dimensions such as objectivity, completeness, accuracy, and representation for research community to validate and compare results.

## Keywords

Conversation maxims, Cooperative principle, Improving information quality, Information quality assessment and analysis, Interrater reliability, Presenting information in relevant way

## 1 Introduction

The international Data Corporation reported that the total amount of global data surpassed 1.8 zettabyte in 2011 and it is predicted to reach 35 zettabyte by 2020 (Kambatla et al., 2014). Economies, companies, and our daily activities are becoming more and more data driven (Zhang et al., 2005; Fidler and Lavbič, 2015). As a result, demand for high-quality information is increasing; however, we are still struggling to understand how quality of information can be improved. Consumers are not interested in just any information; they request the best information available for their purpose (Mai, 2013). This elusive trait of information, how well it serves consumers' needs or how fit it is for use, was investigated further by (Wang and Strong, 1996)

into definition of information quality (IQ) as a multi-dimensional concept with dimensions such as accuracy, consistency, completeness, timeliness, and representation.

The majority of the previous IQ research was conducted by employing students to measure quality of online objects, mostly Wikipedia articles (Kogut and Zander, 1996; Stenmark, 2000). Authors investigated how students perceived IQ dimensions, but did not advance their research with recommendations regarding IQ improvement (Metzger, 2007). A step forward was made by (Mai, 2013), who theoretically discussed the possibility of using the Grice's maxims to construct objects from which consumers could retrieve more relevant information for specific purpose or context of use. Grice's maxims are recommendations used within the cooperative principle for effective communication between two conversing parties, such as be relevant and clear, make messages as informative as possible, and communicate only true and confirmable facts (Grice, 1967). This research aims to experimentally test whether Gricean principle can make Wikipedia articles fit for students' needs, thus improve IQ.

The contribution of this study is to empirically and statistically prove that Gricean principle, if applied to customize information in objects for specific consumers' use, can improve IQ. This study provides evidence of IQ improvement when Wikipedia content is made fit for students' needs with an approach based on Gricean principle. Research community is encouraged to conduct similar IQ improvement studies, not only for students but also for other groups (e.g. young professionals, retired people, etc.). Furthermore, results can be beneficial for Wikipedia and other encyclopedias since they can apply Gricean principle in similar way to improve their content for students' needs.

Additionally, the paper statistically analyzes and interprets students' evaluations of Wikipedia articles by dimensions of accuracy, completeness, representation, and objectivity. This provides the research community with informed insights on students' perception of IQ, which can be used as a reference point in future studies. IQ researchers who employ students as evaluators can validate their results, whereas non-students-based evaluation studies can compare and further generalize their findings.

The remainder of the paper is organized as follows. Section 2 presents a comprehensive review of related work about IQ evaluation and improvement. Section 3 bestows research method for estimating impacts on IQ when cooperative principle is used to make Wikipedia articles more relevant for students' needs. Section 4.2 reports on the results of students' perception of quality with discussion and implication. The paper concludes with Section 5, where reflections on the findings are included and possible directions for future research are discussed.

## 2 IQ

### 2.1 Construct taxonomy

There is no agreed-upon definition of IQ (Michnik and Lo, 2009) and despite significant interest in IQ the domain still remains quite immature (Baskarada and Koronios, 2014). Especially, because the world "quality" characterizes non-physical construct, namely information, which when retrieved and consumed can have different meaning for different users. While some argue that data has only meaning if put into context, thus becoming information (Vrhovec et al., 2015), others emphasize objective and subjective view on information.

Objective view on information is defined as observer and situation independent (Hjørland, 2007) and clarified in (Bates, 2005b) that any (observer and situation independent) difference produces information and therefore also is information. Based on this view, IQ is discussed from external aspect (Arazy and Kopak, 2011) as the degree to which information meets specified and generally accepted requirements (Eppler, 2006). These are usually quality specifications (Ge and Helfert, 2007) set for optimal data values stored in a database (Savchenko, 2003) to avoid deficiencies between the real world state and information system representation (Wand and Wang, 1996).

(Hjørland, 2007) argues that subjective definition of information is at least as or even more important. Subjective aspect (also pointed out as subjectivity) is explained as observer and situation dependent and

illustrated in (Hjørland, 2007) as a difference that makes a difference (for somebody or for something or from a point of view). From this perspective IQ is considered as individual's "subjective judgment of goodness and usefulness" (Hilligoss and Rieh, 2008) or the degree to which the information meets the expectations of the user (Eppler, 2006).

Wang and Strong (1996) adopted "fitness for use" definition, which considers the consumer's perspective, embodying both the objective and subjective perspective of consumed information (Wang and Strong, 1996). Using a two-stage survey, they developed a framework with four categories and 15 dimensions which are generally accepted in the literature (Baskarada and Koronios, 2014). Additionally, other researchers developed a range of IQ evaluation metrics (Madhikermi et al., 2016), assessment instruments (Suhardi et al., 2014) and frameworks (Fink-Shamit and Bar-Ilan, 2008; Hilligoss and Rieh, 2008; Rieh and Danielson, 2007; Bates, 2005a), which can be used to assess different aspects of IQ. (Knight and Burn, 2005) analyzed characteristics of 12 such IQ frameworks, including Wang and Strong's and concluded that each framework is based upon an author's viewpoint, meaning that IQ dimension are combined in different ways (Ge and Helfert, 2007), such as hierarchical (Wang and Strong, 1996), ontological (Wang and Wang, 1996), semiotic (Helfert, 2001), product and service (Kahn et al., 2002).

Recent research emphasized the importance of IQ within wide range of industries, including online communities (Mohammadi et al., 2015; Font et al., 2015; Zheng et al., 2013; Detlor et al., 2013), financial industry (Lee et al., 2016; Corona et al., 2015), healthcare (Lopez et al., 2016; Ceylan et al., 2016), digital media (Romero-Rodriguez et al., 2016), and tourism (Berezan et al., 2016; Paglieri et al., 2014), and demonstrated the business impact that retrieved high-quality information can have on supply chain (Zhou et al., 2014), risk management (Corona et al., 2015; Nicolaou et al., 2013), reporting (Madhikermi et al., 2016), innovativeness, and stock market return (Lee et al., 2016), as well as other impacts, such as positive consumer trust (Berezan et al., 2016), user website satisfaction (Bastida and Huan, 2014), perceived website quality, trust and usefulness (Ghasemaghahi and Hassanein, 2016; Leite et al., 2016), user's decision making (Petter et al., 2013; Shen et al., 2013), customer loyalty, blogging success (Wang et al., 2014), and re-purchase (Ghasemaghahi and Hassanein, 2016). There have been, however, few examinations and validated suggestion on how to improve IQ of objects, due to the opposing perceptions of quality among information consumers.

## 2.2 Students, Wikipedia use and assessment

In recent years Wikipedia received a great interest from research community with the attention on how good is the quality of its articles. Several authors (Denning et al., 2005; Luyt et al., 2008; Wallace and Van Fleet, 2005) expressed concerns about the quality of Wikipedia as a source and interest in evaluation of its content. Overall, people perceived the quality of Wikipedia articles as "quite good" (Chesney, 2006; Stvilia et al., 2008) and often read the articles to obtain additional knowledge (Fallis, 2008). Several other studies indicated (Brown, 2011; Clauson et al., 2008) that Wikipedia users should worry more that Wikipedia articles are incomplete and inaccurate.

There were many initiatives that information-seeking public should be examining and controlling the quality (Ghasemaghahi and Hassanein, 2016; Zheng et al., 2013). As a result, many IQ studies are performed with Wikipedia articles, whereas students are employed as IQ assessors (Mesgari et al., 2015). Despite controversies with student citation of Wikipedia and concerns about poor gatekeeping (e.g. editorial or peer review) (Arazy and Kopak, 2011; Helfert et al., 2013), many researchers do in fact cite Wikipedia and even promote its use (Okoli et al., 2014).

Wikipedia is open for everyone, easily accessible, and interactive (Mai, 2013), and it provides a unique opportunity for educating students in digital literacy (Okoli et al., 2014). In a recent literature review, (Okoli et al., 2014) reported 34 IQ studies in which researches investigated how students use Wikipedia as a general source of information and how they were assigned work that explicitly involved reading Wikipedia articles. (Lim and Kwon, 2010) compared student usage of Wikipedia by gender. They found that while male students used Wikipedia more frequently and had a positive attitude toward it, female students displayed more cautious or conservative attitudes, emotions, and behaviors.

Students like Wikipedia, since it is very comprehensive and easily readable, but at the same time they are well aware of its limitation and they make the best of it (Korosec et al., 2010). (Shaw, 2008) indicated that most students used Wikipedia when they were unfamiliar with the topic as a starting point to obtain basic knowledge that lead to other sources. When students were familiar with the topic, they used Wikipedia to gather additional information about the topic. (Okoli et al., 2014) identified two use cases: personal (source of information) or academic (citations) use. Even when students use Wikipedia for academic purpose they are well aware of its limitations exploiting Wikipedia to find more “reputable” sources. (Choolhun, 2009) for instance documented that Wikipedia is increasingly being used as the first source for legal information inquiries by law students. (Waters, 2007) recommended Wikipedia to students by saying that Wikipedia is a fine place to search for a paper topic or begin the research process. (Patch, 2010) concluded that by employing Wikipedia, students can have an easier time making the leap to higher-level inquiry and responsible scholarship.

## 2.3 Improving IQ

Literature review showed that there is a gap regarding IQ improvement studies. Most of the assessment studies are descriptive, further clarifying taxonomy of IQ by proposing new dimensions. However, they fail to provide actions for IQ improvement of Wikipedia articles. Although it was identified which dimensions assessors consider the most relevant at IQ evaluation, it was at the same time noted that assessors interpret and use dimension values differently. (Arazy and Kopak, 2011) revealed the full extent of this issue, empirically illustrating why past research was not able to identify IQ improvement measures. IQ researchers are therefore challenged to find alternative approaches to identify IQ improvement measures and prove the significance of their impact.

Although cooperative principle with conversation maxims arise from pragmatics of natural language and are used to improve conversational effectiveness of communication, (Mai, 2013) proposed and discussed their use in the field of IQ. The author elaborated that IQ must be understood in a context in which the consumer is situated while retrieving information. Information producers should therefore act as if they are actually speaking with information consumers and overcome the disordered nature of a language by applying Gricean principle (Grice, 1967) for successful communication referred to as maxims and submaxims.

Studied literature review showed that there are no empirical studies, which investigated the application of cooperative principle to improve IQ. It is assumed that presenting information in a more relevant way for the specific use improves the perception of IQ, but it is yet not known to what extent and by what rate of agreement among consumers. Thus, knowing why users’ need information provides one with the capability for questioning poor IQ and suggesting new solutions for its improvement.

This research narrows focus on investigation of IQ improvement of Wikipedia articles when Gricean principle is applied. Therefore, main research question is as follows:

**RQ1.** Can Grice’s maxims and submaxims be applied to Wikipedia articles to improve their quality for students?

In particular, this paper investigates what the overall IQ affect rate is and whether there are some recognized IQ dimensions that are more affected than others.

## 3 Method

To test the effect of cooperative principle on Wikipedia articles, a quantitative study with 265 students of computer information studies at entry level (35 percent female, 65 percent male; from 20 to 26 years of age, with the mean of 21,90) was performed. Each student was invited to complete an online survey active from January 2015 until December 2017, where they were asked to evaluate three selected and re-written articles from Wikipedia, presented to them in a random order by four IQ dimensions: accuracy, completeness, objectivity, and representation as presented in Table 1. Selected IQ dimensions represented most relevant

Table 1: Selected IQ dimensions

| Dimension      | Description  |
|----------------|--|
| Accuracy       | Information in the article is accurate.  |
| Completeness   | The article is complete and includes all necessary information.                  |
| Objectivity    | The article is objective; it represents objective opinion about presented topic. |
| Representation | The article is presented consistently and formatted concisely.                   |

top-level IQ categories (intrinsic, contextual, and representational) and were comparable to dimensions used in other research (Sackmann, 1991; Arazy and Kopak, 2011).

In January 2015 three original Wikipedia articles with different topics: human (homo-sapiens from “People and self” Wikipedia category), alexandrite (gemstone with color changing ability from “Math science and technology” Wikipedia category) and Occitan language (Romance language mostly spoken in Southern France from “Culture and arts” Wikipedia category) were selected. Participants had extensive domain knowledge about human article opposed to narrow domain knowledge about the alexandrite and Occitan language article. Articles from selected topics established similar domain knowledge distribution across participants’ as in recent study (Arazy and Kopak, 2011).

Grice’s maxims and submaxims were used to re-write original Wikipedia articles and make them more relevant for student cognitive needs, identified and described in (Lim, 2009). Gricean principle was applied in such a manner that information was presented to participants in the same, standardized way, serving both needs, balancing general with detailed information. This avoided possible user preference toward characteristics of object which was used to present the information, e.g. the length of an article abstract, the number of included sources, the writing style, and the reputation of included references (Knight and Burn, 2005).

In the online survey 83 percent of students replied to be annoyed by low IQ while 16 percent were not annoyed by it. They were asked to evaluate each re-written Wikipedia article by all four IQ dimensions with Likert scale of 1 (strongly disagree) to 7 (strongly agree). All three articles were presented to students in a random order as well as the order of evaluated dimensions was randomized per article to avoid confounding bias (Pannucci and Wilkins, 2010). Collected assessments were analyzed with statistical methods: mean, standard deviation, Cronbach’s  $\alpha$ , and ICC (two-way random model of single measures). In order to answer our research questions, we compared the results of our student subgroup with the recent study of (Arazy and Kopak, 2011) and discussed results.

### 3.1 Proposed approach of applying Gricean principle

Cooperative principle was used to shorten original Wikipedia articles (Mai, 2013) and made them more relevant for typical Wikipedia users, namely students. (Lim, 2009) reported that students use Wikipedia to satisfy their cognitive needs, such as: look for quick facts, to learn something that they are not familiar with and to get more information about familiar topics. Therefore, original Wikipedia articles, obtained on January 2015, were summarized and structured in paragraphs combining general information about most important entity facts in detail, according to Grice’s principles of conversation maxims and submaxims from Table 2.

Articles summaries were structured with five consecutive paragraphs, each of two to three sentences long. In the first paragraph entity used in the article was described. Next, it was explained in more details how this particular entity differs from similar or related entities. Then two paragraphs were used to present basic information about entity’s history (explained origin, historic development and status of the entity) and to explain basic characteristics of the entity, such as average size, color, quantity, etc. In the last two paragraphs, the basic information about entities’ impact on society and a couple of interesting, specific facts (trivia) associated with entity were added.

In total, each re-written article had four paragraphs (1, 3, 4, 5) and a picture (how entity looks) to present

Table 2: Grice’s conversation maxims and submaxims used to re-write Wikipedia articles

| Category                         | Maxims and submaxims  | Applied action   |
|----------------------------------|---|--|
| C. Relation                      | C1. Be relevant   | Articles summaries were structured with consecutive 5 paragraphs: definition (explained object, discussed in article), differentiation (explained differences from similar or related objects), history (explained origin, historic development and current status of the object), basic characteristics (explained characteristics of the objects such as average size, color, quantity, etc.) and social impact (explained impact on society). |
| A. Quantity                      | A1. Make your contribution as informative as is required (for the current purposes of exchange) | Article summaries were made as informative as possible within structured paragraphs.   |
|                                  | A2. Do not make your contribution more informative than is required                             | Two pictures were included in the article.   |
| B. Quality                       | B1. Do not say what you believe to be false   | Summarized information was cross-checked by various other sources, such as encyclopedias (Britannica, Scholarpedia, Encarta, etc.) and specialized sites for chosen topics.  |
|                                  | B2. Do not say that for which you lack adequate evidence  |  |
| D. Manner<br>(be<br>perspicuous) | D1. Avoid obscurity of expression   | Concise and orderly writing style was used.  |
|                                  | D2. Avoid ambiguity   |  |
|                                  | D3. Be brief (avoid unnecessary prolixity)  |  |
|                                  | D4. Be orderly  |  |

general information about the entity and two paragraphs (2 and 6) and a picture to present more detailed information. Reference pages, quotations or other sources used in the original Wikipedia article were skipped to avoid potential biases or preferences of participants toward the, e.g., credibility of some sources. Final objects, re-written Wikipedia article, used in the research had title, summary (1 - 6 paragraphs), and two pictures as shown in Figure 1. This approach of object construction significantly reduces the amount of used words in re-written articles compared to Wikipedia originals. For instance, the article about Human when newly constructed had only 180 words compared to 12.000 words of the same original article on Wikipedia.

### 3.2 Data analysis

Information quality  $IQ^{(dim)}$  for a given dimension  $dim$  (accuracy, completeness, representation, and objectivity) was defined as a mean value of all marks given by  $k$  raters and is given as follows:

$$IQ^{(dim)} = \frac{1}{k} \sum_{i=1}^k x_i^{(dim)}$$

where  $x_i$  refers to a specific rater score.

In order to calculate agreement level for all four IQ dimensions, interclass correlation ( $ICC(2, 1)$ , agreement in two-way model with single measures) was employed and defined as:

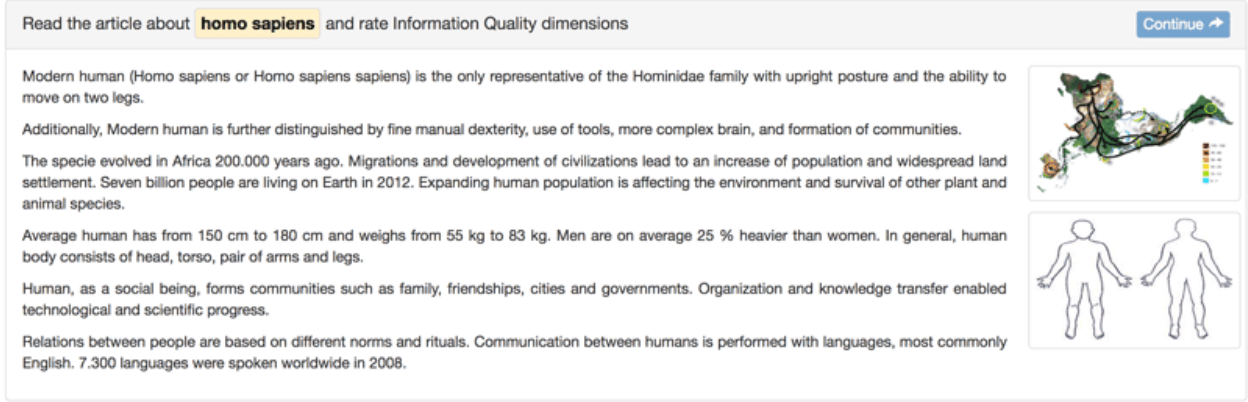


Figure 1: Outlook of re-written human Wikipedia article with applied Gricean principle

$$\frac{var(\beta)}{var(\alpha) + var(\beta) + var(\epsilon)}$$

where  $var(\beta) = \frac{BMS-WMS}{k}$  is a variability due to differences in the objects,  $var(\alpha) = \frac{JMS-EMS}{n}$  is a variability due to differences in ratings levels used by raters and  $var(\epsilon) = EMS$  is a variability due to differences in the evaluations of the objects by the raters.  $BMS$  is the subject mean square,  $WMS$  is the residual mean square,  $JMS$  is the rater mean square, and  $EMS$  is the object/rater mean square.

When calculating agreement level for specific IQ dimensions normalized standard deviation  $AL_{nsd}^{(dim)} = 1 - \frac{s^{(dim)}}{s_{max}^{(dim)}}$  was used, where standard deviation  $s^{(dim)}$  is defined as  $\sqrt{\frac{(x_i^{(dim)} - \bar{x}^{(dim)})^2}{n-1}}$  and maximum standard deviation  $s_{max}^{(dim)}$  on a seven-point Likert scale is defined as  $3 \cdot \sqrt{\frac{n}{n-1}}$ . A similar approach for calculating agreement level was employed in (Haakonsen Dahl and Jørgensen, 2014; Moe-Nilssen et al., 2008), where authors relied on absolute reliability and the smallest detectable difference.

## 4 Results and discussion

In order to validate and analytically investigate if Wikipedia articles, which were made fit for students with applied Gricean principle, had improved IQ, this paper's results (**FLR**) were compared with Arazy and Kopak research (**AKR**) (Arazy and Kopak, 2011). Arazy and Kopak used the same IQ dimensions, scale, and user group (student in early twenties) to measure IQ. In Section 4.1 evidence of IQ improvement is supported with results and three key findings. Results about student's perception of individual IQ dimensions are presented in Section 4.2, followed by a discussion of the obtained findings in Section 4.3.

### 4.1 Evidence of IQ improvement

Results of article evaluations are presented as key statistical IQ indicators in Table 3 and compared with results published in the AKR. The average value of scores for all four IQ dimensions (CIQ) increased by 0,27 compared to the AKR. At the same time reliability of construct was improved by 0,18 (Cronbach's  $\alpha$ ) and consistency among raters increased by 0,03 ( $ICC(2,1)$ ). The values of Cronbach's  $\alpha$  were well above the required 0,75 threshold (Straub et al., 2004) and  $ICC(2,1)$  had a  $p$ -value lower than the 5% significance level. These indicators confirmed that students in the FLR sample perceived IQ of re-written Wikipedia articles to be higher than students who evaluated IQ of original Wikipedia articles in the AKR.

Table 3: Overall research results compared to AKR

| Indicator                       | FLR    | AKR    | diff  |
|---------------------------------|--------|--------|-------|
| Number of students              | 265,00 | 270,00 | -5,00 |
| Reliability Cronbach's $\alpha$ | 0,98   | 0,80   | 0,18  |
| Consistency $ICC(2,1)$          | 0,20   | 0,17   | 0,03  |
| CIQ mean value                  | 5,16   | 4,89   | 0,27  |

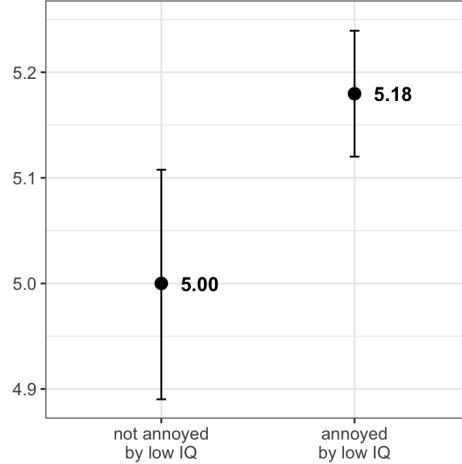


Figure 2: CIQ means with confidence intervals for students who answered not to be annoyed by low IQ and those who were annoyed

Additionally, difference in perception of CIQ was also investigated internally, only for students who participated in FLR research, by dividing them in two groups: those who answered to be annoyed by low IQ (83%) and those who were not (17%).

Difference in evaluation of CIQ between both groups  $\overline{IQ}_{\text{not annoyed by low IQ}}^{(CIQ)} = 5,00 < \overline{IQ}_{\text{annoyed by low IQ}}^{(CIQ)} = 5,18$  is presented in Figure 2 and is statistically significant (Mann-Whitney-Wilcoxon test;  $\alpha = 0,05$ ) with  $W = 6,32 \times 10^{-5}$  and  $p = 0,0153$ . This indicates that students who considered themselves more sensitive to low IQ recognized application of Gricean principle as an improvement. Furthermore, even students who were not annoyed by low IQ considered Wikipedia article re-written with Gricean principle still to have higher CIQ (5,00) than original Wikipedia article in the AKR (4,89).

## 4.2 Results of IQ dimensions evaluations

Impact on quality (mean value of scores) and agreement level (normalized standard deviation of scores) of making Wikipedia content fit for students' needs is further presented by each IQ dimension in Figure 3. AKR reported accuracy, objectivity, and representation to be similar in size regarding absolute mean and agreement level values. Completeness on the other hand had low values of both indicators. Making Wikipedia articles more relevant to students' needs changed the proportions of these measures. In terms of IQ mean value three levels were formed: high, which was achieved by objectivity; moderate, obtained by representation; and low, achieved by accuracy and completeness. The agreement level of accuracy, completeness, and representation was higher as of completeness.

Impacts per individual IQ dimensions are depicted in Figure 4, where assessment results are compared to the AKR in even more straight forward manner. The IQ score of completeness increased the most by 1,01 (on 1 – 7 scale), followed by 0,67 increase of objectivity, while the IQ score of representation declined by



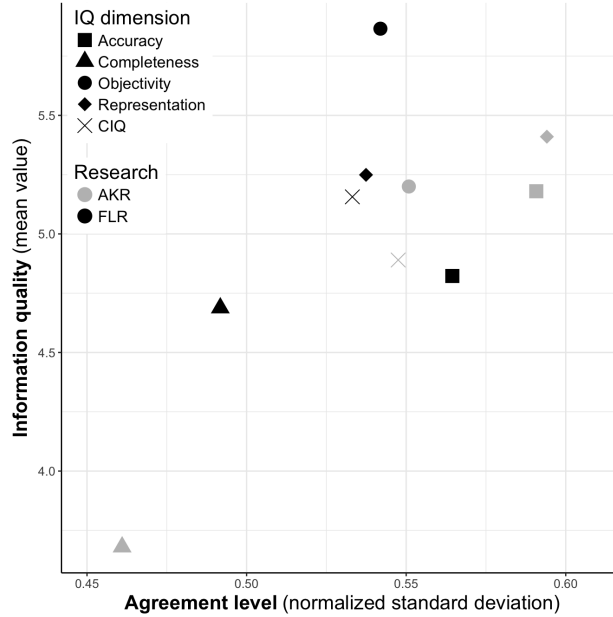


Figure 3: Impact of individual IQ dimensions on IQ and agreement, compared to AKR

0,16 and accuracy by 0,36. In terms of agreement level, the normalized standard deviation of completeness slightly increased by 0,03, while it marginally declined for dimensions of objectivity (by 0,01), accuracy (by 0,03), and Representation (by 0,06).

### 4.3 Discussion of impact on IQ dimensions

The IQ score of accuracy decreased due to polarized participant’s domain knowledge distribution across selected articles. Students considered their domain knowledge about the evaluated article either expert or ignorant. In all, 71% of the students rated the article of Occitan and Alexandrite as 4, not being able to agree or disagree regarding the article’s quality. If such assessments are discarded, value for accuracy actually increases by 0,77 points.

Approach of applied Gricean principle which made Wikipedia articles fit for students needs also increased their perception of quality for dimensions of objectivity and completeness. Students perceived retrieved information from modified Wikipedia articles more objective and complete as students retrieved information from randomly chosen original Wikipedia articles that were evaluated in the AKR. Thus, the agreement level between students’ scores for completeness increased, making a dimension more measurable, while agreement level of objectivity stayed the same.

The quality of objectivity most likely increased because Grice’s conversation maxims and submaxims (Grice, 1967) were applied with very succinct writing style that was perceived by students as more objective as the writing style of Wikipedia editors. In addition, all given information in re-written articles was cross-checked with other online sources causing also the perception of accuracy to increase, but only for students who were confident enough (had enough domain knowledge) to evaluate this dimension. Both improvements made re-written articles less biased and more truthful.

Even more noticeable finding is the improvement of completeness. As discussed in (Yaari et al., 2011), article length is not entirely detached from the content. Authors argue that longer articles, if written comprehensively, include more information than shorter ones. In other words it would be difficult to write short articles of good quality that would still be considered complete. Although application of Gricean principle made re-written articles only 180 words long, students perceived shorter versions to be more complete, compared to randomly

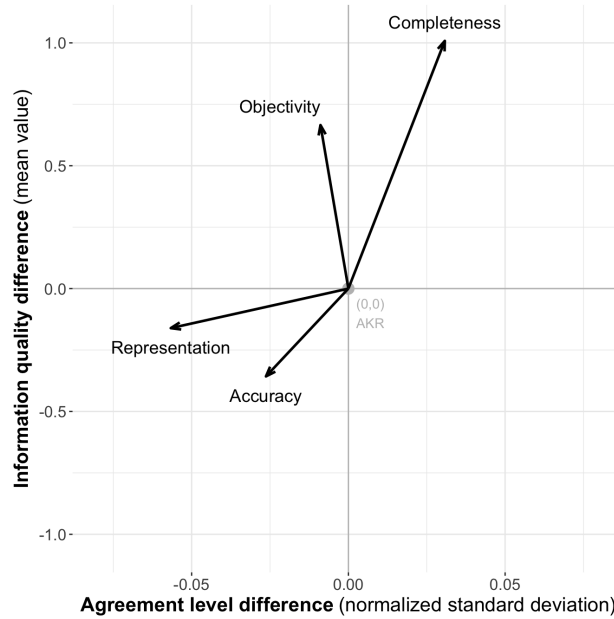


Figure 4: Difference of agreement level and IQ for individual dimensions as vectors compared to AKR (0, 0)

selected articles from Wikipedia with length from 200 to 3.500 words, which were evaluated in the AKR. There are two implications of this finding.

First, encyclopedias could implement Grice’s conversation maxims and submaxims as editorial guidelines or use proposed structure of re-written articles as a template for presenting concise summaries of topics. This study confirmed that there are many possibilities for IQ improvement and that editing based on Gricean principle would make articles for students more complete and objective. Further more, presenting information in more concise manner makes articles fitter for students since they typically use Wikipedia as a starting point for further inquiries. Second, researchers could gain more IQ evaluations if they would shorten their articles. Rater spends only one-fifth of the time to read an article of 180 words as opposed to an article of 1.000 words. Researchers could therefore use each rater to assess more articles or offer less time demanding surveys.

Despite, providing visually appealing assessment form and consistent presentation of information in articles (e.g. same notation for number and currencies), students did not perceive quality of representation to improve compared to the AKR. It may be that students are not accustomed to assessing representation to the same extent as for other IQ dimensions. In addition, information representation is not so often critical for daily tasks and even when it becomes critical, it can be more easily resolved as inaccurate or incomplete information.

## 5 Conclusion

This paper proposed and tested an approach of making Wikipedia articles more relevant for students’ needs, thus provided evidence of improved IQ. (Mai, 2013) expectations that the pragmatic philosophy of a language can be employed to make objects more relevant for consumer needs were confirmed. When Gricean principle was applied on Wikipedia articles to make them fit for students’ needs the average value of scores for all four IQ dimensions (CIQ) increased by 0,27 compared to the AKR (Arazy and Kopak, 2011). Students annoyed by low IQ recognized re-written articles to have significantly higher IQ (for 0, 18) than students who were not annoyed by low IQ. IQ increase was notable for dimensions of objectivity and completeness, quality of accuracy decreased due to polarized domain knowledge and representation obtained similar quality as in the AKR.

This paper provides to authors informed insights on students' perception of IQ, which can be used as a reference point in future studies. IQ researchers who employ students as evaluators can validate their results, whereas non-students-based evaluation studies can compare and further generalize their findings. Furthermore, results can be beneficial for Wikipedia and other encyclopedias since they can apply Gricean principle in a similar way to improve their content for students' needs. In this manner Wikipedia articles can be shortened, serving students with information of improved IQ but at the same time demanding less work for editors. Students can read more comprehensive articles faster, obtaining information of a higher quality in less time, while editors would maintain articles of fewer words.

This study has limitations, since it was based on one application of Grice's conversation maxims and submaxims; on a limited number of articles; compared to a limited number of recent researches; and conducted for students. Despite our consulting and academic background, which promotes succinct written and verbal communication and literature review on research how and why students use Wikipedia, it is not conclusive that proposed application of Grice's conversation maxims and submaxims was the most effective communication of relevant information to students. IQ of Wikipedia articles can be improved even more. Only limited number of articles were selected. However, they were selected in such a manner that students evaluated topics of low and high domain knowledge. Results were compared only to the AKR, since there is a deficiency of similar evaluation studies.

Research community is therefore encouraged to conduct similar IQ improvement studies, not only for students but also for other groups (e.g. young professionals, retired people etc.). Proposed approach of applying Gricean principle can be used as a recommendation on how to construct articles of fewer words to have shorter questionnaires, obtain more article evaluations per rater, and ultimately have more reliable measurements. In the future we hope to repeat our study with a larger sample size of articles of more diverse raters, based on characteristics such as age, education, and domain knowledge, so that findings could be generalized to entire population with higher confidence.

## References

- Arazy, O. and Kopak, R. (2011). On the Measurability of Information Quality. *Journal of the American Society for Information Science and Technology*, 62(1):89–99.
- Baskarada, S. and Koronios, A. (2014). A Critical Success Factor Framework for Information Quality Management. *Information Systems Management*, 31(4):276–295.
- Bastida, U. and Huan, T. C. (2014). Performance evaluation of tourism websites' information quality of four global destination brands: Beijing, Hong Kong, Shanghai, and Taipei. *Journal of Business Research*, 67(2):167–170.
- Bates, M. J. (2005a). Information and knowledge: an evolutionary framework for information science. *Information Research*, 10(4).
- Bates, M. J. (2005b). An Introduction to Metatheories, Theories, and Models. In *Theories of Information Behaviour*, pages 1–24.
- Berezan, O., Yoo, M., and Christodoulidou, N. (2016). The impact of communication channels on communication style and information quality for hotel loyalty programs. *Journal of Hospitality and Tourism Technology*, 7(1):100–116.
- Brown, A. R. (2011). Wikipedia as a Data Source for Political Scientists: Accuracy and Completeness of Coverage. *PS - Political Science & Politics*, 44(2):339–343.
- Ceylan, H. H., Caypinar, B., Kucukkoc, M., Uzuner, S., and Kucukdurmaz, F. (2016). Information Quality on Developmental Dysplasia of the Hip on Turkish Websites. *Journal of Academic Research in Medicine-Jarem*, 6(2):84–87.
- Chesney, T. (2006). An empirical examination of Wikipedia's credibility. *First monday*, 11(11).

- Choolhun, N. (2009). Google: to use, or not to use. What is the question? *Legal Information Management*, 9(3):168–172.
- Clauson, K. A., Polen, H. H., Boulos, M. N. K., and Dzenowagis, J. H. (2008). Scope, Completeness, and Accuracy of Drug Information in Wikipedia. *Annals of Pharmacotherapy*, 42(12):1814–1821.
- Corona, C., Nan, L., and Zhang, G. Q. (2015). Accounting Information Quality, Interbank Competition, and Bank Risk-Taking. *Accounting Review*, 90(3):967–985.
- Denning, P., Horning, J., Parnas, D., and Weinstein, L. (2005). Wikipedia risks. *Communications of the ACM*, 48(12):152–152.
- Detlor, B., Hupfer, M. E., Ruhi, U., and Zhao, L. (2013). Information quality and community municipal portal use. *Government Information Quarterly*, 30(1):23–32.
- Eppler, M. J. (2006). *Managing Information Quality: Increasing the Value of Information in Knowledge-intensive Products and Processes*. Springer-Verlag, Berlin, Germany, 2nd edition.
- Fallis, D. (2008). Toward an epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology*, 59(10):1662–1674.
- Fidler, M. and Lavbič, D. (2015). Research About Measurability of Information Quality. In *10th International Conference on Knowledge Management in Organisations (KMO 2015)*, Slovenia, Maribor. 2015/KMO.
- Fink-Shamit, N. and Bar-Ilan, J. (2008). Information quality assessment on the Web - an expression of behaviour. *Information Research - an International Electronic Journal*, 13(4).
- Font, L., Zouaq, A., and Gagnon, M. (2015). *Assessing the Quality of Domain Concepts Descriptions in DBpedia*. 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems. Ieee, New York.
- Ge, M. and Helfert, M. (2007). *A Review of Information Quality Research - Develop a Research Agenda*.
- Ghasemaghaei, M. and Hassanein, K. (2016). A macro model of online information quality perceptions: A review and synthesis of the literature. *Computers in Human Behavior*, 55:972–991.
- Grice, H. P. (1967). *Logic and Conversation*. Harvard University Press.
- Haakonsen Dahl, S. S. and Jørgensen, L. (2014). Intra- and Inter-Rater Reliability of the Mini-Balance Evaluation Systems Test in Individuals with Stroke. *International Journal of Physical Medicine & Rehabilitation*, 2(1).
- Helfert, M. (2001). *Managing and Measuring Data Quality in Data Warehousing*.
- Helfert, M., Walshe, R., and Gurrin, C. (2013). The Impact of Information Quality on Quality of Life: An Information Quality Oriented Framework. *Ieice Transactions on Communications*, E96B(2):404–409.
- Hillgoss, B. and Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, 44(4):1467–1484.
- Hjørland, B. (2007). Information: Objective or subjective/situational? *Journal of the American Society for Information Science and Technology*, 58(10):1448–1456.
- Kahn, B. K., Strong, D. M., and Wang, R. Y. (2002). Information Quality Benchmarks: Product and Service Performance. *Communications of the ACM*, 45(4).
- Kambatla, K., Kollias, G., Kumar, V., and Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, 74(7):2561–2573.
- Knight, S.-A. and Burn, J. (2005). Developing a Framework for Assessing Information Quality on the World Wide Web. *Informing Science Journal*, 8:159–172.

- Kogut, B. and Zander, U. (1996). What firms do? Coordination, identity, and learning. *Organization Science*, 7(5):502–518.
- Korosec, L., Limacher, P. A., Lüthi, H. P., and Brändle, M. P. (2010). Chemical information media in the chemistry lecture hall: a comparative assessment of two online encyclopedias. *CHIMIA International Journal for Chemistry*, 64(5):309–314.
- Lee, R. P., Chen, Q. M., and Hartmann, N. N. (2016). Enhancing Stock Market Return with New Product Preannouncements: The Role of Information Quality and Innovativeness. *Journal of Product Innovation Management*, 33(4):455–471.
- Leite, P., Goncalves, J., Teixeira, P., and Rocha, A. (2016). A model for the evaluation of data quality in health unit websites. *Health Informatics Journal*, 22(3):479–495.
- Lim, S. (2009). How and Why Do College Students Use Wikipedia? *Journal of the American Society for Information Science and Technology*, 60(11):2189–2202.
- Lim, S. and Kwon, N. (2010). Gender differences in information behavior concerning Wikipedia, an unorthodox information source? *Library & Information Science Research*, 32(3):212–220.
- Lopez, D. M., Blobel, B., and Gonzalez, C. (2016). Information quality in healthcare social media - an architectural approach. *Health and Technology*, 6(1):17–25.
- Luyt, B., Aaron, T. C. H., Thian, L. H., and Hong, C. K. (2008). Improving Wikipedia’s accuracy: Is edit age a solution? *Journal of the American Society for Information Science and Technology*, 59(2):318–330.
- Madhikermi, M., Kubler, S., Robert, J., Buda, A., and Framling, K. (2016). Data quality assessment of maintenance reporting procedures. *Expert Systems with Applications*, 63:145–164.
- Mai, J. E. (2013). The quality and qualities of information. *Journal of the American Society for Information Science and Technology*, 64(4):675–688.
- Mesgari, M., Okoli, C., Mehdi, M., Nielsen, F. A., and Lanamaki, A. (2015). "The Sum of All Human Knowledge": A Systematic Review of Scholarly Research on the Content of Wikipedia. *Journal of the Association for Information Science and Technology*, 66(2):219–245.
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13):2078–2091.
- Michnik, J. and Lo, M. C. (2009). The assessment of the information quality with the aid of multiple criteria analysis. *European Journal of Operational Research*, 195(3):850–856.
- Moe-Nilssen, R., Nordin, E., and Lundin-Olsson, L. (2008). Criteria for evaluation of measurement properties of clinical balance measures for use in fall prevention studies. *Journal of Evaluation in Clinical Practice*, 14(2):236–240.
- Mohammadi, F., Abrizah, A., and Nazari, M. (2015). Is the information fit for use? Exploring teachers perceived information quality indicators for Farsi web-based learning resources. *Malaysian Journal of Library & Information Science*, 20(1):99–122.
- Nicolaou, A. I., Ibrahim, M., and van Heck, E. (2013). Information quality, trust, and risk perceptions in electronic data exchanges. *Decision Support Systems*, 54(2):986–996.
- Okoli, C., Mesgari, M., Mehdi, M., Nielsen, F. A., and Lanamaki, A. (2014). Wikipedia in the Eyes of Its Beholders: A Systematic Review of Scholarly Research on Wikipedia Readers and Readership. *Journal of the Association for Information Science and Technology*, 65(12):2381–2403.
- Paglieri, F., Castelfranchi, C., Pereira, C. D., Falcone, R., Tettamanzi, A., and Villata, S. (2014). Trusting the messenger because of the message: feedback dynamics from information quality to source evaluation. *Computational and Mathematical Organization Theory*, 20(2):176–194.

- Pannucci, C. J. and Wilkins, E. G. (2010). Identifying and Avoiding Bias in Research. *Plastic and Reconstructive Surgery*, 126(2):619–625.
- Patch, P. (2010). Meeting Student Writers Where They Are: Using Wikipedia to Teach Responsible Scholarship. *Teaching English in the Two-Year College*, 37(3):278–285.
- Petter, S., delone, W., and McLean, E. R. (2013). Information Systems Success: The Quest for the Independent Variables. *Journal of Management Information Systems*, 29(4):7–61.
- Rieh, S. Y. and Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, 41:307–364.
- Romero-Rodriguez, L. M., de Casas-Moreno, P., and Torres-Toukourmidis, A. (2016). Dimensions and Indicators of the Information Quality in Digital Media. *Comunicar*, (49):91–100.
- Sackmann, S. A. (1991). *Cultural Knowledge in Organizations: Exploring the Collective Mind*. SAGE Publications.
- Savchenko, S. (2003). Automating Objective Data Quality Assessment (Experiences in Software Tool Design). In *International Conference on Information Quality (IQ 2003)*.
- Shaw, D. (2008). Wikipedia in the Newsroom. *American Journalism Review*, pages 40–45.
- Shen, X. L., Wang, N., Sun, Y. Q., and Xiang, L. (2013). Unleash the power of mobile word-of-mouth: An empirical study of system and information characteristics in ubiquitous decision making. *Online Information Review*, 37(1):42–60.
- Stenmark, D. (2000). Leveraging tacit organizational knowledge. *Journal of Management Information Systems*, 17(3):9–24.
- Straub, D., Boudreau, M.-C., and Gefen, D. (2004). Validation guidelines for IS positivist research. *The Communications of the Association for Information Systems*, 13(1):380–427.
- Stvilia, B., Twidale, M. B., Smith, L. C., and Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and Technology*, 59(6):983–1001.
- Suhardi, Gunawan, G., Dewi, A. Y., and Ieee (2014). Total Information Quality Management-Capability Maturity Model (TIQM-CMM): An Information Quality Management Maturity Model. *2014 International Conference on Data and Software Engineering (ICODSE)*, page 6.
- Vrhovec, S. L. R., Trkman, M., Kumer, A., Krisper, M., and Vavpotic, D. (2015). Outsourcing as an Economic Development Tool in Transition Economies: Scattered Global Software Development. *Information Technology for Development*, 21(3):445–459.
- Wallace, D. P. and Van Fleet, C. (2005). The democratization of information? Wikipedia as a reference resource. *Reference & User Services Quarterly*, 45(2):100–103.
- Wand, Y. and Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11):86–95.
- Wang, R. Y. and Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, 12(4):5–33.
- Wang, Y. S., Li, H. T., Li, C. R., and Wang, C. (2014). A model for assessing blog-based learning systems success. *Online Information Review*, 38(7):969–990.
- Waters, N. L. (2007). Why you can’t cite Wikipedia in my class. *Communications of the ACM*, 50(9):15–17.
- Yaari, E., Baruchson-Arbib, S., and Bar-Ilan, J. (2011). Information quality assessment of community-generated content - A user study of Wikipedia. *Journal of Information Science*, 37(5):487–498.

- Zhang, X. M., Anghelescu, H. G. B., and Yuan, X. J. (2005). Domain knowledge, search behaviour, and search effectiveness of engineering and science students: an exploratory study. *Information Research - An International Electronic Journal*, 10(2).
- Zheng, Y. M., Zhao, K. X., and Stylianou, A. (2013). The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation. *Decision Support Systems*, 56:513–524.
- Zhou, H. G., Shou, Y. Y., Zhai, X., Li, L., Wood, C., and Wu, X. B. (2014). Supply chain practice and information quality: A supply chain strategy study. *International Journal of Production Economics*, 147:624–633.