Altmetrics as Additional Indicators for Analyzing Research Efficiency of Universities? A Conceptual and Empirical Study

By Dr. Dirk Tunger, TH Köln und Forschungszentrum Jülich GmbH, Prof. Dr. Heinz Ahn, Technische Universität Braunschweig, Prof. Dr. Marcel Clermont, Duale Hochschule Gera-Eisenach, Dipl.-Volksw. and Dipl.-Kff. Johanna Krolak, Forschungszentrum Jülich GmbH, Julia Langner, M.Sc., Technische Universität Braunschweig, and Dr. Bernhard Mittermaier, Forschungszentrum Jülich GmbH*)

The academic literature provides a variety of empirical studies measuring the research efficiency of universities. The indicators used are mainly bibliometric values, which are intended to quantify the dissemination and impact of university research in the scientific community. Further objectives of a university's management related to research activities that refer to other stakeholders such as the interested public are typically not addressed in these studies. Against this background, the question arises whether altmetrics are suitable for quantifying such neglected objectives in order to be used as additional indicators in multidimensional efficiency analyses. Based on a framework built on fundamentals of decision theory, the paper conceptually addresses this question by discussing the potential of altmetrics to quantify research objectives. It turns out that only few altmetrics actually meet essential measurement requirements. The remaining altmetrics are then used in an empirical study to investigate their impact on research efficiency of three different research fields by means of Data Envelopment Analysis. It is shown that the inclusion of altmetrics, ceteris paribus, does not lead to any significant changes in the ranking positions of the universities under consideration.

1 Introduction

Universities are under increased pressure concerning research due to competition, especially for rare publication space and third-party funding. It is, therefore, necessary for a university's management to efficiently allocate the available resources. However, evaluating the research efficiency of universities as a whole or of their

*) Dr. *Dirk Tunger* is team lead of a third-party funded project dealing with Altmetrics at TH Köln, Faculty of Information and Communication Sciences, Institute for Information Science. Prof. Dr. *Heinz Ahn* is head of the Institute of Management Control & Business Accounting at Technische Universität Braunschweig. Prof. Dr. *Marcel Clermont* is professor for Financial and Management Accounting and head of Retail Management at Cooperative University Gera-Eisenach. Dipl.-Volksw. & Dipl.-Kff. *Johanna Krolak* was research assistant at the Central Library at Forschungszentrum Jülich GmbH. *Julia Langner*, M.Sc., is a research assistant at the Institute of Management Control & Business Accounting at Technische Universität Braunschweig. Dr. *Bernhard Mittermaier* is head of the Central Library at Forschungszentrum Jülich GmbH. Email: d.tunger@fz-juelich.de, hw.ahn@tu-braunschweig.de, marcel.clermont@dhge.de, bonnjoanna@aol.de, j.langner@tu-braunschweig.de, b.mittermaier@fz-juelich.de. Funding: This paper is based on two projects funded by the German Federal Ministry of Education and Research under grant numbers 16PU17001 and 16IFI107. The responsibility for the content lies with the authors.

organizational units (e. g., faculties) entails major challenges. In light of these challenges, there is a series of different approaches to gain insights by measuring research efficiency based on performance indicators, especially bibliometric measures. In the literature, there are several case studies concerning the analysis of academic research efficiency.[1] To date, however, no standard approach has been established.

Criticism regarding such studies primarily relates to the choice of performance indicators. It is often noted that the indicators are selected/constructed more or less arbitrarily or that the set of indicators is incomplete.[2] It is, therefore, necessary, as a first step, to base efficiency measurement on a theoretical foundation. One approach in this regard is provided by the theory of (prescriptive) decision-making originating from business administration. The main advantage of this theory is its systematic focus on the objectives pursued. In this approach, indicators are measures for quantifying these objectives. Based on this, coordinated steps can be derived for an efficiency analysis, which are intended to counteract arbitrariness in the selection and construction of indicators. Therefore, we will first introduce a suitable decision-oriented framework in the next chapter in order to be able to systematically embed our performance analysis outlined below.

The focus of our paper is the inclusion of universities' digital communication in traditional analyses of research efficiency. Like in many areas of private and business life, more and more processes, results, and discussions in science are shifting to the digital world.[3] Scientific output, e. g., a journal article, is shared and discussed on established social media such as Twitter and Facebook. There are also more and more platforms created specifically for scientists, such as Academia.edu, ResearchGate, and Mendeley. Against this background, we explore the question of what possibilities and limits exist to date for taking social media activities of higher education institutions into account in research efficiency analyses.

While the objective of distributing research findings is traditionally quantified by bibliometric indicators such as the number of publications and citations, the distribution in digital media requires complementary indicators. Web-based metrics, which are summarized under the generic term "alternative metrics" or "altmetrics", are in principle suitable for this purpose.[4] With regard to the use of altmetrics in research evaluations, there is currently widespread skepticism in the literature. For example, *Bornmann/Haunschild* (2016a, 2016b) mention that a normalization of indicators on a field level is not possible at the moment or that data can be manipulated through a practice known as gaming.[5]

However, to date, there is a lack of a systematic inclusion of web-based indicators in research efficiency analyses. In this paper, we want to pursue an initial conceptual approach in this respect. Based on concepts of decision-making theory (Chapter 2), we conceptually analyze which altmetrics are capable of mapping research objectives that may be pursued from a university management point of view (Chapter 3). We demonstrate these considerations by means of empirical examples to analyze the informative value especially of the altmetrics that are taken into account (Chapter 4).

1 See *Rhaiem* (2017) for an overview.
2 See *Clermont/Dirksen* (2015).
3 See *Tunger/Clermont/Meier* (2018).
4 See *Priem/Taraborelli/Groth/Neylon* (2010).
5 See *Biagioli/Lippman* (2020).

The contribution of our paper consists of three core aspects:

- Development of a decision-oriented framework for analyzing research efficiency from the perspective of university management
- Conceptual analysis of the applicability of altmetrics with regard to measuring research efficiency, resulting in three principally suitable indicators
- Initial empirical efficiency analysis for different research fields of universities, showing that the three chosen altmetrics have currently little impact on the efficiency results stemming from traditional research performance indicators

2 A Decision-oriented Framework for Performance Measurement

Performance measurement as a field of research in business administration deals with how the performance of organizational units can be ascertained and evaluated. This research field has been rather an ad-hoc science for a long time. While the instruments that have been developed (e. g., Balanced Scorecard) address a problem of business practice, many of them lack a theoretical basis (they are also referred to as management methods). This deficit can also be observed with regard to performance measurement for universities. Economists sometimes use indicators from the research field of bibliometrics, such as the number of publications and citations, to evaluate the research performance of individual researchers or departments with little further consideration.

At the end of the 1990s, the first approaches to theoretically ground performance measurement emerged. From a business administration point of view, an expedient approach is to link considerations of performance measurement with decision theory. In this context, multi-criteria production theory was developed by *Dyckhoff* (2006, 2018). In this generalized production theory, the focus of the traditional approach on the process of object transformation is basically retained. Production is understood as a qualitative, quantitative, spatial, or temporal change of objects, which is initiated by people and systematically directed and carried out to create value.[6] While conventional production theory solely relates to input/output quantities and their cost/revenue, multi-criteria production theory also allows for the systematic consideration of other objectives of a transformation process. This is achieved by including the decision-makers (manufacturer, production manager, or other kinds of stakeholder, hereinafter referred to as decision-making unit (DMU)) and their preferences in the analysis.

The corresponding evaluation of a production process encompasses five components:[7]

- Alternatives: these represent the realizable production possibilities – or more generally, the realizable activities – to choose from
- Consequences: these specify principally all effects resulting from the realization of the chosen alternative in a neutral manner
- Uncertainties: being beyond the control of the decision-maker, these cause the alternatives' consequences to vary depending on the conditions that occur after the decision has been made

6 See *Dyckhoff* (2006), p. 3.
7 See *Eisenführ/Weber/Langer* (2010), p. 20.

- Objectives: these determine which of the consequences the decision-maker is interested in – in a positive or negative sense
- Preferences: these express the utility with which the decision-maker associates the consequences, e. g., relative to each other and in terms of their value, occurrence in time, and risk perception.

The alternatives, their consequences, and the uncertainties are considered to be objective in nature. By mapping them with the individual objectives and preferences of the DMU in a specific situation, the alternatives can be evaluated. A series of theory-based approaches have been developed for this purpose. Ideally, they condense the data of a multi-objective decision-making problem into a single utility score that allows the alternatives to be compared.

Of course, it is not the primary purpose of performance measurement to support decisions. However, its purpose is to evaluate their success with respect to the responsible DMU. Since the decisions, as outlined above, have been made based on the DMU's objectives, the achievement of these objectives is the obvious measure of success. In non-trivial cases, objective-related indicators are used to quantify these achievements. If several objectives are to be taken into account using indicators with different scales, the challenge of how to aggregate these indicators into a one-dimensional score arises. This is often done by merging them into a measure of effectiveness or efficiency, called effectiveness/efficiency degree.

An effectiveness degree quantifies the extent to which a DMU achieved the intended purposes of decisions, i. e. the primary objectives. An efficiency degree also includes the resources needed, which can be interpreted as objectives to be minimized; accordingly, the achieved primary objectives are weighed up against the consumed resources. In addition, efficiency must always be determined in relation to the DMUs being assessed. Therefore, it is referred to as a relative efficiency measurement.[8] A DMU is considered efficient if no other unit exists that is at least as good with regard to all objectives and better with regard to at least one of them.[9]

A performance analysis should be conducted as part of a comprehensible procedure in order to increase the decision quality and guarantee intersubjective verifiability. Building on the aforementioned theoretical foundations, *Ahn/Clermont* (2018) propose a corresponding guideline based on existing, more specific procedures.[10] The authors differentiate between a planning phase (P) and a measuring phase (M), with each comprising three steps. They suggest starting with defining the boundaries of the performance analysis (step P1); the target group of the analysis is key to this. Once the boundaries have been determined, the target group's objectives are revealed (step P2). As these objectives are often insufficient in operational terms, it has to be determined which indicators should be used to measure the extent to which the objectives are met (step P3). This task involves the consideration of several requirements. From a pragmatic point of view, data availability and the economic viability of data collection have to be considered. In qualitative terms, the indicators' validity and reliability are of crucial relevance.[11]

8 See *Clermont* (2016), p. 1353.
9 See *Ahn/Le* (2016).
10 See *Emrouznejad/Witte* (2010).
11 See *Kromrey/Roose/Strübing* (2016), pp. 167–168/242–245.

To actually measure performance, it is necessary to consider the options for data collection and to perform a preliminary analysis of the collected data (step M1). By aggregating the different objective-related indicators, an overall effectiveness or efficiency degree of the assessed DMUs can then be quantified; there is a range of methods available to do so (step M2). Finally, the results have to be evaluated (step M3).

In the following, we will use the described framework to analyze the applicability of altmetrics as additional indicators for evaluating the research efficiency of universities from the perspective of their management (step P1). Chapter 3 outlines which objectives university managers could pursue with research by including web-based aspects (step P2) and addresses the questions of which indicators are capable of depicting these web-based aspects (step P3). In Chapter 4, an exemplary efficiency analysis is carried out by means of Data Envelopment Analysis (phase M).

3 Research Objectives and Altmetrics as Indicators

3.1 Research-related Objectives of Universities

According to *Van Vught/Ziegele* (2010), universities represent multi-purpose organizations, i. e., they pursue several (overall) objectives at the same time. In this context, it should be noted that universities do not have a uniform structure both internationally and within certain countries, accompanied by different task priorities. This makes it necessary to adopt a certain stakeholder perspective in order to derive a system of objectives for universities. Here, we will focus on publicly funded universities in Germany.

A first source of information on relevant objectives is the German Higher Education Framework Act ("Hochschulrahmengesetz"), which specifies the normative framework for universities in Germany. Another source are agreements between the federal states and their universities on objectives to be achieved and services to be provided.[12] They are often formulated as part of a benchmarking process. Finally, other stakeholders, like third-party funders and the general public, are becoming more prominent in higher education policy, thus expanding the list of objectives.[13] From a synopsis of these three sources, the following research-related objectives can be derived, which are relevant for the analysis of research efficiency from the perspective of university management with respect to different stakeholder groups:

- Research represents a main task of universities. Its underlying objective is the generation of knowledge and its dissemination in the professional scientific community.[14]
- A second objective associated with research is to maintain/increase competitiveness, i. e., outstanding research contributes to an attractive environment for scientists, managing staff, and students, among others.[15]
- Another objective gaining attention is to take civic responsibility. This manifests itself, e. g., in practical research with economic and social benefits. Corresponding opportunities are manifold and range from promoting the region to volun-

12 See *Jongbloed/De Boer/Kolster/Kottmann/Vossensteyn/Benneworth/Cremonini/Lemens-Krug* (2015).
13 See *In der Smitten/Jaeger* (2012).
14 See *Chalmers* (1999).
15 See *Cartalos* (2012).

tary commitments and international engagement. In this regard, some universities face increasing pressure to legitimize their research efforts in order to secure acceptance from policymakers, special interest groups, and the interested public at large.[16] Other universities consider this as a chance to engage in social media activities as part of their strategy.

3.2 *Indicators for Measuring Research-related Objectives*

The indicators used in economics to measure research efficiency relate to the first two objectives listed above. Commonly, these are bibliometric indicators, such as the (weighted) number of publications and citations. They primarily address the generation of knowledge and its dissemination in the professional scientific community, but may also serve as proxies for a university's competitiveness resulting from outstanding research.[17] Another proxy for both of these objectives are third-party funds that are also often investigated in research efficiency studies. Such funds are assumed to reflect past high quality research results and also enable, e. g., capital-intensive research activities, creating an attractive environment for scientists and students interested in science.[18]

The mentioned bibliometric indicators and third-party funds cannot address the (facets of the) third objective properly, i. e., taking civic responsibility to legitimize a university's research efforts from the perspective of the public at large and other interest groups concentrated on a special subject. To measure this objective, a fruitful data source is offered by the impact indicators that emerged from the new forms of communication via the World Wide Web, since they reflect other effects, such as social, educational, and cultural ones. These indicators were developed within the framework of the scientometric sub-disciplines of altmetrics. With respect to the collection of altmetric data, users can meanwhile rely on professional data providers,[19] like altmetric.com. There, the following altmetrics can be retrieved based on the DOI of an article: Blog posts, F1000 posts, Facebook posts, Mendeley readership, News stories, Patents, Peer reviews, Policy documents, Q&A posts, Reddit posts, Tweets, Videos, and Wikipedia pages.

In order to be used both as indicators for the quantitative mapping of the outlined objectives and as indicators for an efficiency analysis, altmetrics must fulfill construct-related and application-related requirements.[20] The former include validity and reliability. An indicator is valid if it is suitable for measuring the facts that are of actual interest, i. e., the degree to which an objective is attained. It is reliable if repeating the measurement under the same conditions produces the same result. The application-related requirements include the availability of a sufficiently large amount of data that also allows discrimination between the DMUs considered.

Especially due to the application-related requirements, many of the altmetrics collected by altmetric.com cannot be considered for an efficiency analysis. In a preliminary investigation, we have found out that Blog posts, F1000 posts, Facebook posts, Patents, Peer reviews, Policy documents, Q&A posts, Reddit posts, Videos, and Wiki-

16 See *Melo/Figueiredo* (2019).
17 See *Clermont/Dirksen/Scheidt/Tunger* (2017).
18 See *Melo/Figueiredo* (2019).
19 See *Bar-Ilan/Halevi/Milojević* (2019).
20 See *Kromrey/Roose/Strübing* (2016), pp. 179–180; *Rassenhövel* (2010), p. 23.

pedia pages have both a low number of mentions and no discriminatory power. Therefore, only Mendeley readership, News stories, and Tweets are considered in the following.

Mendeley represents a reference management program that can be used to upload, organize, exchange, and cite scientific publications. It also functions as a social network, allowing users to communicate with other scientists. The altmetric indicator used here is the number of article downloads, called Mendeley readership.[21] Mendeley is mainly used by scientists, i. e., by the professional scientific community so that activities in the context of accessing Mendeley reflect the attention of corresponding users to a topic. This is particularly evident in the fact that studies have found high correlations between Mendeley readership and citations.[22]

It is concluded that the number of Mendeley readership and citations measure similar things. This is not surprising, since both indicators can be traced back to activities of the same group of stakeholders, the scientists themselves. However, there is a discrepancy in terms of time: while research papers can be added to Mendeley and then downloaded immediately without any preconditions being met, there is a time delay on citations within classic scientific communication, e. g., caused by the often protracted nature of the peer review process.[23] Accordingly, *Costas/Zahedi/Wouters* (2015), as well as *Mohammadi/Thelwall/Kousha* (2016) show that the number of Mendeley readership is a suitable early indicator of citations. The number of Mendeley readership thus appears valid for measuring the objective of disseminating research results.[24] However, there are limits on the reliability since no information is available as to when a publication was downloaded – the Mendeley readership cannot be precisely determined for a period under analysis. In addition, the number of Mendeley readership might decline after a certain date of analysis because of accounts being deleted. Therefore, repeated measurements may generate different results.[25] In terms of application-related requirements, Mendeley readership sufficiently provides data and allows for differentiation between the units to be evaluated.

News stories are texts of any length that aim to raise awareness of a topic. With reference to the discussion of scientific publications, it can be assumed that News stories are predominantly written by science journalists for the general public that has an interest in science. News stories indicate that the underlying publications have sufficient potential to attract public attention.[26] They thus represent, in principle, a valid indicator for mapping the objective of civic responsibility. However, its reliability is limited, since News stories are predominantly published on websites on the World Wide Web, which can be deleted or deactivated. This means that the number of News stories over a particular survey period can produce different results at different times. With respect to application-related requirements, News stories provide a sufficiently large amount of data, which also enables to generate differentiating analysis results.

Tweets comprise a short message of not more than 280 characters (until 2017: 140 characters) in length on the microblogging service Twitter. Other modes of communication linked to these Tweets also exist in the form of retweets, tags, likes, refer-

21 See *Mohammadi/Thelwall/Kousha* (2016).
22 See, e. g., *Thelwall* (2018); *Breuer/Schaer/Tunger* (2020).
23 See *Mittermaier* (2020).
24 See *Van Noorden* (2014).
25 See *Zahedi/Costas/Wouters* (2017).
26 See *Kousha/Thelwall* (2019).

ences to other user profiles, following other user profiles, links, and images. There are several analyses that show low correlations between Tweets and citation indicators.[27] Based on this, it can be concluded that Tweets measure aspects that are different from citations and that Twitter is primarily used by stakeholders outside of science. Accordingly, it is supposed that Tweets relating to research publications mainly illustrate the social impact, the speed at which knowledge is disseminated or transferred, and public interest in a particular subject.[28] Thus, at least from a conceptual point of view, they appear to be valid for mapping the objective of civic responsibility. A restrictive statement must be made regarding reliability since the number of Tweets measured in a period under examination is not stable. For example, Tweets and Twitter accounts can be deleted. Furthermore, Tweets originally shared to the general public can later be set to be visible only to a particular group of people, or vice versa. As a result, it cannot be guaranteed that the same results are obtained if a measurement is repeated later on. With regard to the application-related requirements, a large amount of data can be generated, which also allows a sufficiently good differentiation of results.

4 Effects of Altmetrics in Efficiency Analyses

In the following, we will analyze the impact of altmetrics on the efficiency determination of universities based on empirical data. As previously outlined, there are numerous studies on the correlation of altmetrics with classic bibliometric indicators. However, to the best of our knowledge, their impact on efficiency analyses has not yet been determined. For this reason, our study is exploratory in its design: our aim is to gain fundamental insights and uncover problems associated with measuring efficiency based on altmetrics. These findings might then be used in subsequent studies to develop theories founded on hypotheses. For our empirical example, we use Data Envelopment Analysis (DEA), which is a well-established approach, often used for measuring the efficiency of universities.[29] Below, we briefly introduce the DEA concept before discussing our data set and the results of our efficiency analyses.

4.1 Measuring Efficiency Using Data Envelopment Analysis

DEA is a non-parametric method because it does not require specifying the particular form of the production function in advance. Furthermore, it is deterministic, meaning that the distance of a DMU from the production function is only attributed to the DMU's inefficiency. There are two classic approaches to quantify such a distance: output orientation and input orientation. They differ in how the inefficient DMUs are projected onto the production function. In the case of output orientation, the factor is determined by which all assessed outputs have to be increased proportionally to reach the production function while the inputs remain constant for the assessed DMU. In contrast, input orientation results in the necessary proportional reduction of the inputs while the outputs remain constant. In the context assessed here – just as in the literature for the most part[30] – output orientation is used since it is easier for universities to influence their outputs than their inputs.

27 See *Haustein/Peters/Bar-Ilan/Priem/Shema/Terliesner* (2014); *Ortega* (2016).
28 See *Eysenbach* (2011).
29 See *Rhaiem* (2017); *Ahn/Clermont/Langner* (2021).
30 See *Rhaiem* (2017).

The basic models of DEA differ in terms of the returns to scale that are assumed for the efficient production processes. While the CCR model, named after its founders *Charnes, Cooper, and Rhodes*, is based on constant returns to scale,[31] the BCC model proposed by *Banker, Charnes, and Cooper* is based on variable returns to scale.[32] The production technology and production function of the BCC model result directly from the assumptions of non-emptiness, free disposability, and convexity, while the CCR model also implies scalability of each input-output combination.

In the literature, variable returns to scale are overwhelmingly assumed.[33] We adhere to this assumption and therefore present the mathematical formulation of the BCC model. There are two approaches to this formulation: the envelopment form and the multiplier form. The envelopment form refers to the production technology being formed through the envelopment of the given input-output data. In contrast, the multiplier form puts inputs and outputs into proportion and assigns relative weights – so-called multipliers – to them endogenously. Since the conventional DEA models are based on linear optimization, the model forms can be transferred to each other in line with the duality theorem.[34] We use both forms hereinafter, as they produce analysis results with differences in detail.

Let us assume that $x \in \mathbb{R}_+^m$ represents m semi-positive inputs of a DMU j, which are transformed into non-negative outputs, $y \in \mathbb{R}_+^s$. To determine the efficiency of a DMU j in the envelopment form, the optimization problem

$$\begin{aligned} (Eff_j)^{-1} = \phi_j^* = max\ \phi_j & \\ s.t.\ \ & X\lambda \leq x_j \\ & Y\lambda \geq \phi_j y_j \\ & \underline{1}^T\lambda = 1 \\ & \lambda \geq 0 \end{aligned} \tag{1}$$

must be solved. This identifies the highest proportional multiplication ϕ_j of the outputs that is still within the production technology. All DMUs h with $\lambda_h > 0$ are reference DMUs. They determine the benchmark $(X\lambda, Y\lambda)$ of an inefficient DMU. In this respect, λ_h describes the respective proportion of DMU h that forms the benchmark.

The multiplier form puts the inputs and outputs into proportion based on endogenously determined weights. To this end, the following optimization problem rendered as a linear-fractional equation must be solved:

$$\begin{aligned} (Eff_j)^{-1} = \phi_j^* = min\ \frac{ux_j+w}{vy_j} & \\ s.t.\ \ & \frac{uX+w}{vY} \geq 1 \\ & u \geq 0 \\ & v \geq 0 \\ & w \in \mathbb{R}. \end{aligned} \tag{2}$$

The parameters u_i and v_r can be interpreted as implicit relative weights of a marginal input unit i and output unit r, respectively. They represent the contribution of input and output factors to determine efficiency and are referred to as multipliers.

31 See *Charnes/Cooper/Rhodes* (1978).
32 See *Banker/Charnes/Cooper* (1984).
33 See *Clermont/Dirksen/Dyckhoff* (2015).
34 See *Charnes/Cooper/Rhodes* (1978).

The relative weights are specifically set for each DMU so that every DMU appears as good as possible. It must be noted that these are no absolute weights. Thus, they can be compared with each other only in relative terms by taking the order of magnitude into account, for example through standardization.

4.2 Data Set and Descriptive Analysis

It must be determined which DMUs (i. e., which universities) are examined. Since altmetrics, as other bibliometric indicators, have different distributions across research fields, we performed the analyses for each research field individually after conducting a pretest. The goal of this pretest was to determine three research fields that distinguish themselves by a high, a satisfactory, and a poor number of the altmetrics chosen by us. As a result, hereinafter we assess the fields of medicine (high number), engineering (satisfactory number), and economics (poor number).

The inputs and outputs to be considered must also be determined. In this respect, personnel and financial variables are usually used as inputs, while publications are almost always used as outputs, occasionally complemented by citations.[35] We will also use such classic indicators hereinafter to remain consistent with the literature and avoid misinterpretations. Thus, the number of professors and the number of scientific employees as well as the expenditures on research equipment in the respective field at the university in question are used as inputs. We obtained respective data from the German Federal Statistical Office. Journal publications and citations of these publications, again broken down by field and university, are used as classic outputs. These variables were gathered using the Web of Science database hosted by the Competence Centre for Bibliometrics.[36] The publications were allocated to research fields based on the journal in which the article was published. The science classification by *Archambault/Beauchesne/Caruso* (2011) was used to avoid overlap in the allocation of the journals to fields. This 3-level classification system allocates each journal to one subfield, which can be aggregated at higher levels to form fields and domains.

The DOIs of the publication sets generated in this way were then used to gather the three selected altmetrics in order to ensure that the data refer to the same publications with an identical allocation to a specific research field. Our data was collected on 16 March 2020 by using altmetric.com.

A period of five years (2014–2018), over which all the indicators were aggregated, was selected for the analyses. This allows potential fluctuations to be balanced and a potential delay in the reaction to publications to be anticipated. The universities and their respective fields that are studied here are those for which all data for all years was available. This amounts to 44 universities in medicine, 49 universities in engineering, and 83 universities in economics. The publications are allocated at the university level through disambiguation of the affiliations of scientific institutions in Germany. As we intend to evaluate the suitability of altmetrics for measuring the

35 See *Ahn/Clermont/Langner* (2021).

36 In order to allow the scientific community to perform comprehensive and reliable bibliometric analyses, the Federal Ministry of Education and Research (BMBF) launched the German Competence Centre for Bibliometrics (grant No. 16WIK2101A) in 2008. It has the aim of providing adequate data infrastructure to perform scientifically demanding and complex bibliometric analyses. Further details are available at http://www.bibliometrie.info/en.

efficiency of universities in general, there is deliberately no assessment of individual universities, and no rankings are given.

Table 1 gives descriptive data on the inputs and outputs for medicine, engineering, and economics. The data reflect the intended substantial difference between the research fields. It is also apparent that the altmetrics are skewed to the right. Their values are low for most universities while being particularly high for a few others. This is also the case for publications and citations.

		Professors	Research assistants	Expenditures (in mio €)	Publications	Citations	Mendeley readership	News stories	Tweets
Medicine	Min	10	13	0.38	94	593	1912	18	245
	Max	1154	19021	8739	13991	192003	514964	7899	117881
	Mean	431	6381	2287	4314	50420	140538	2192	28751
	Median	465	7411	2211	3992	45218	132283	1928	26226
	Std. dev.	273	4839	2073	3130	41098	112954	1735	25001
	Skewness	0.43	0.45	0.86	1.20	1.40	1.50	1.50	1.90
Engineering	Min	6	0	0.44	3	14	150	1	9
	Max	1010	12515	576	4933	57068	177503	3058	35295
	Mean	262	2440	93	1134	12485	32267	656	6828
	Median	198	1437	51	814	7841	17478	487	4091
	Std. dev.	238	2914	115	1190	13706	37537	684	7914
	Skewness	1.05	1.71	2.06	1.80	1.70	1.80	1.50	1.70
Economics	Min	12	0	0.57	1	12	82	0	0
	Max	451	1609	330	1283	5218	40274	550	5951
	Mean	128	445	15	166	776	5356	79	681
	Median	113	376	6	131	580	3833	40	412
	Std. dev.	78	327	38	169	746	5628	110	839
	Skewness	1.74	1.27	7.22	3.90	3.00	3.40	2.70	3.60

Table 1: Descriptive data on the inputs and outputs[37]

4.3 *Efficiency Analyses with One Output*

We have used the aforementioned inputs and outputs to perform efficiency analyses using DEA with variable returns to scale. To this end, we first compared the three inputs to one output for each research field separately. Five outputs were considered, the three altmetrics discussed in detail in Chapter 3 (Mendeley readership, News stories, and Tweets) as well as publications and citations. This resulted in a total of 15 individual analyses (three research fields times five outputs).

37 The values given are rounded and based on the respective total of the variable values for 2014–2018.

		Publications	Citations	Mendeley readership	News stories	Tweets
Medicine	Publications	1	0.96	0.98	0.94	0.88
	Citations	0.96	1	0.98	0.93	0.85
	Mendeley readership	0.98	0.98	1	0.95	0.89
	News stories	0.94	0.93	0.95	1	0.90
	Tweets	0.88	0.85	0.89	0.90	1
Engineering	Publications	1	0.96	0.91	0.91	0.92
	Citations	0.96	1	0.97	0.93	0.96
	Mendeley readership	0.91	0.97	1	0.90	0.95
	News stories	0.91	0.93	0.90	1	0.94
	Tweets	0.92	0.96	0.95	0.94	1
Economics	Publications	1	0.89	0.84	0.69	0.86
	Citations	0.89	1	0.93	0.70	0.82
	Mendeley readership	0.84	0.93	1	0.70	0.82
	News stories	0.69	0.70	0.70	1	0.82
	Tweets	0.86	0.82	0.82	0.82	1

All results are highly significant with $p < 0.001$.

Table 2: Rank correlation of pairs of outputs

To identify more easily what fundamental relationships exist between the assessed indicators, we first calculated their Spearman's rank correlations, listed in Table 2. If the rank correlation is 1, the assessed rankings are identical. As it can be seen from Table 2, all rank correlations are consistently high. The highest values for Tweets refer to engineering, while the highest values for the other indicators refer to medicine. As described before, the Mendeley readership is seen as an early indicator for citations, which seems to be confirmed by the correspondingly very high rank correlations of both criteria.

Table 3 shows the Spearman rank correlations of the DEA efficiencies including the stated outputs compared to the DEA efficiencies including publications as output. Again, all results are again highly significant, and a clear difference between the research fields can be seen. However, in the results, engineering has the highest rank correlations, especially with respect to Tweets and Mendeley readership. This is not in line with the results from Table 2. Thus, we cannot conclude that a high (rank) correlation of the altmetrics themselves implies a high rank correlation of the respective efficiency results. In general, there is not necessarily a strong connection between (rank) correlation in data and rank correlation of efficiency results, since DEA is not translation invariant.[38]

38 See *Dyson/Allen/Camanho/Podinovski/Sarrico/Shale* (2001).

	Output	Spearman rank correlation
Medicine	Citations	0.935
	Mendeley readership	0.935
	News stories	0.852
	Tweets	0.793
Engineering	Citations	0.976
	Mendeley readership	0.945
	News stories	0.936
	Tweets	0.971
Economics	Citations	0.811
	Mendeley readership	0.735
	News stories	0.454
	Tweets	0.745

Table 3: Rank correlation of the efficiencies based on publications versus altmetrics

The three altmetrics considered have sufficient discriminatory power for the research fields and period under assessment. However, what happens if we use an indicator as an output that does not guarantee this? To answer this question, we analyzed exemplarily the efficiency using the Q&A posts for all fields, while otherwise keeping the conditions constant. Now, the rank correlation with the efficiencies based on publications is just 2.6 % in economics, for example. This result could signify that Q&A posts actually depict different objectives than publications. However, such a conclusion cannot be based solely on low rank correlations, which would in fact be a misinterpretation in the present case. Instead, they can be attributed to the fact that in economics, for example, over 90 % of the assessed universities achieve an efficiency of 0 % because they have not published any Q&A posts. Thus, the results of this approach are of scarce significance.

4.4 Efficiency Analysis with Two Outputs

As already outlined, there is no expectation at this time that altmetrics can or should replace traditional indicators for performance. Instead, they are designed to provide additional information and, above all, to provide information regarding the perception of scientific publications more quickly than it is possible using bibliometrics. Therefore, considering altmetrics as an additional output should be reviewed. To this end, we conducted analyses considering the publications together with citations and, alternatively, with each of the three altmetrics previously addressed. Once again, we have determined rank correlations for the resulting efficiencies, which are listed in Table 4. In addition, data about the normalized multipliers that result from the multiplier model are listed for each of the two included indicators. As explained in Chapter 4.1, multipliers have the function of relatively weighting the indicators to provide information on the relevance of including the respective indicator in the efficiency determination. The normalized multipliers can be interpreted as percentage weights.

The rank correlations are very similar to the results from Table 3. Once again, engineering has the highest rank correlations. However, the values in Table 4 are based on efficiency analyses in which one of the two outputs (publications) is identical. As a result, this inevitably produces higher rank correlations among the efficiencies. The relevance of the publications appears in the multipliers. The normalized multipliers referring to the publications have a significantly higher mean than the normalized multipliers referring to the other output in each case. The median of the latter mostly equals 0, meaning that only for a few universities the second output does have a positive effect on the efficiency determined.

Second output		Rank correlation	Normalized multiplier for publications				Normalized multiplier for the second output			
			Mean	Median	Std. dev.	Skewness	Mean	Median	Std. dev.	Skewness
Medicine	Citations	1	0.84	1	0.35	-1.99	0.15	0	0.35	1.99
	Mendeley readership	0.993	0.71	1	0.44	-0.98	0.29	0	0.45	1.00
	News stories	0.968	0.72	1	0.41	-0.98	0.28	0	0.41	0.98
	Tweets	0.961	0.85	1	0.34	-2.12	0.15	0	0.34	2.11
Engineering	Citations	1	0.79	1	0.40	-1.41	0.21	0	0.40	1.41
	Mendeley readership	0.997	0.77	1	0.36	-1.38	0.23	0	0.36	1.36
	News stories	0.974	0.66	0.86	0.36	-0.31	0.34	0.14	0.36	0.31
	Tweets	0.997	0.85	1	0.27	-1.93	0.15	0	0.27	1.93
Economics	Citations	1	0.70	1	0.40	-0.84	0.30	0	0.40	0.84
	Mendeley readership	0.970	0.69	1	0.45	-0.87	0.31	0	0.45	0.87
	News stories	0.792	0.70	1	0.40	-0.83	0.30	0	0.40	0.83
	Tweets	0.869	0.70	1	0.45	-0.87	0.30	0	0.45	0.87

The results are highly significant with $p < 0.001$.

Table 4: Rank correlation of the efficiencies and data about the multipliers in the case of combining publications with one other output[39]

It should be noted that a high rank correlation is not necessarily desirable. Including an indicator in analyses, which produces efficiencies that are highly correlated with the efficiencies resulting from classic indicators, provides no additional benefit. Therefore, particularly in medicine and engineering, for the efficiency analyses, virtually no new information can be gained by including altmetrics. Even the relatively low rank correlations of News stories and Tweets in economics are not associated with any additional information content, since they can be attributed to the low indicator level of economics on altmetric.com. Only the efficiencies of a few DMUs benefit from applying this additional indicator. This applies, in particular, to the Mendeley readership. They do, however, provide an advantage insofar as they act as an early indicator of citations.

39 The minimum and maximum are not indicated here because they are 0 and 1, respectively, for all indicators.

Finally, we also examine what happens if an altmetric indicator with low discriminatory power is also included in the efficiency analysis. To this end, we in turn assess the Q&A posts and perform a DEA analysis for each research field with publications as an additional output. The low discriminatory power causes the Q&A posts to have barely any effect on the efficiencies of universities. Thus, the median weight of this indicator is 0 and the mean is very low at 5 % for economics. However, a small number of universities also benefit, since they have several Q&A posts but few publications. In these instances, just a few posts are enough to allocate a very high efficiency of up to 100 % to the universities in question.

5 Summary and Outlook

A review of the state of the art shows that currently no altmetrics are used as indicators in efficiency analyses. In fact, the primary focus of research is on the generation of research findings and their dissemination in the respective scientific community, normally quantified by publication and citation figures. The impact of research outside this scientific community did not play a considerable role in the past. Currently, this is changing as the scientific community is increasingly working on direct interaction with society. Altmetrics offer a path to close this gap, as they make scientific communication visible in web-based media. However, our conceptual considerations reveal that only a few altmetrics currently meet the application-oriented requirements for an inclusion in performance analyses, especially due to the small amount of data and insufficient discriminatory power. Therefore, we have identified only three altmetrics to be included in our empirical analyses: Mendeley readership, News stories, and Tweets.

We used empirical data to analyze the effect of these altmetrics on efficiency results for different research fields. It has become apparent that if altmetrics are included in analyses in addition to publications, the altmetric indicators only have a slight impact on the ranking of universities. This is even the case for medicine, a research field with a fairly good amount of data. Hence, it makes actual no major difference in the resulting rankings if altmetrics are included or not. One advantage, of course, is that altmetrics are available much earlier in the research cycle and are also much more multifaceted.

Our investigations represent initial reflections on the possibilities and limits of the use of altmetrics as indicators in efficiency analyses. Thus, numerous possibilities for further research result from our considerations. For example, it makes sense to systematically derive a system of research objectives from the perspective of university management. Such a system of objectives consists of overall objectives and subobjectives, which have to fulfill basic decision-theoretical requirements. However, the development of a corresponding system of research objectives is still pending, since this is a challenging task and the result has to be validated by empirical studies.

Concerning our chosen three altmetrics, there exists extensive criticism in the literature. For example, the inclusion of a publication in a scientist's Mendeley collection does not necessarily express appreciation (or criticism), as it would be the case if a scientist cites this publication.[40] The number of Mendeley readership merely indicates a general interest in the publications' topic. Furthermore, Mendeley readership

40 See *Garfield* (1996).

only includes the activity of one of the largest reference management systems, which is strongly linked to Scopus users. The number of News stories indicates the extent to which activities of universities were perceived by the interested public. Without a detailed content analysis, however, it is not possible to determine whether the News stories refer to activities relating to civic responsibility. Additionally, while it can be commented on News stories, it remains unclear to what extent such comments should be taken into consideration. Content analyses of Tweets suggest that Tweets tend to reflect a non-specific interest in articles rather than approval, use, or engagement with the research behind them.[41] This question can also only be answered on the basis of detailed content analyses. In addition, there is criticism of easy manipulation regarding the number of Tweets, and the messages often lack substance due to their shortness, or banality.

The future will reveal whether recognized indicators can be deduced from altmetrics for quantifying university management objectives and for using them in efficiency analyses. Further deliberations are needed in any case for a normalization mechanism with respect to different research fields. Initial approaches to their standardization already exist,[42] but they are still far from being mature. It has also been shown that there are high rank correlations between publications with high cognitive relevance scores, high citation counts, and high altmetrics counts, which are rooted in relevance theory.[43]

Finally, it should be mentioned that we are aware that both bibliometric indicators and altmetrics have not been developed for being used in performance measurement analyses of scientists or research institutions.[44] However, the empirical reality is different, in which bibliometric indicators in particular are used in ranking processes as a proxy for the performance and impact of researchers and universities. From this point of view, altmetrics, as explained, in principle offer the possibility of quantifying university managements' objectives that have not been addressed so far in performance analyses. Notwithstanding the above, there is always the risk of disincentives, which can hardly be avoided with any type of performance measurement. A scientist may only try to increase the indicator expression itself, e. g., within a Twitter network, without contributing to the intended objective, i. e., interesting a broad public for the research topic. Furthermore, especially in the case of multiple tasks, scientists may focus on easily measurable achievements, disregarding the ones that are difficult to measure. However, the consideration of multidimensional objectives and corresponding indicators can mitigate such effects. There are initial conceptual analyses of disincentives through rankings and efficiency analyses in higher education,[45] but empirical evidence is still lacking, so there is a need for corresponding research here, especially concerning long-term comparative analyses.

41 See *Holmberg/Thelwall* (2014); *Robinson-Garcia/Costas/Isett/Melkers/Hicks* (2017).
42 See *Bornmann/Haunschild* (2016a, 2016b).
43 See *Breuer/Schaer/Tunger* (2020).
44 See *Tunger/Meier/Hartmann* (2017).
45 See *Osteloh/Frey* (2014).

References

Ahn, H./Clermont, M. (2018), Performance Measurement, in: *Corsten, H./Gössinger, R./Spengler, T. S.* (Eds.), Handbuch Produktions- und Logistikmanagement in Wertschöpfungsnetzwerken, de Gruyter, pp. 886–903.

Ahn, H./Clermont, M./Langner J. (2021), Effizienzmessung in der Hochschule: Verfahrenseinsatz und Implikationen für das Hochschulmanagement, Hochschulmanagement, Vol. 16, No. 1, pp. 4–13.

Ahn, H./Le, M. H. (2016), Decision-oriented performance measurement framework: An emphasis on rationality, in: *Ahn, H./Clermont, M./Souren, R.* (Eds.), Nachhaltiges Entscheiden: Beiträge zum multiperspektivischen Performancemanagement von Wertschöpfungsprozessen, Springer Gabler, pp. 369–383.

Archambault, É./Beauchesne, O. H./Caruso, J. (2011), Towards a multilingual, comprehensive and open scientific journal ontology, in: *Noyons, E./Ngulube, P./Leta, J.* (Eds.), Proceedings of the 13th International Society of Scientometrics and Infometrics Conference, Leiden University and University of Zululand, pp. 66–77.

Banker, R. D./Charnes, A./Cooper, W. W. (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, Management Science, Vol. 30, No. 9, pp. 1078–1092.

Bar-Ilan, J./Halevi, G./Milojević, S. (2019), Differences between altmetric data sources: A case study, Journal of Altmetrics, Vol. 2, No. 1, p. 1.

Biagioli M./Lippman A. (2020), Gaming the metrics: Misconduct and manipulation in academic research, Cambridge.

Bornmann, L./Haunschild, R. (2016a), How to normalize Twitter counts? A first attempt based on journals in the Twitter Index, Scientometrics, Vol. 107, No. 3, pp. 1405–1422.

Bornmann, L./Haunschild, R. (2016b), Normalization of Mendeley reader counts for impact assessment, Journal of Informetrics, Vol. 10, No. 1, pp. 62–73.

Breuer, T./Schaer, P./Tunger, D. (2020), Relations between relevance assessments, bibliometrics and altmetrics, in: *Cabanac, G./Frommholz, I./Mayr, P.* (Eds.), Proceedings of the 10th International Workshop on Bibliometric-enhanced Information Retrieval, pp. 101–112.

Cartalos,A.(2012),TheattractivenessoftheEUfortopscientists,in:https://www.europarl.europa.eu/document/activities/cont/201209/20120904ATT50415/20120904ATT50415EN.pdf, retrieved on 15.11.2021.

Chalmers, A. F. (1999), Wege der Wissenschaft, Springer.

Charnes, A./Cooper, W. W./Rhodes, E. (1978), Measuring the efficiency of decision making units, European Journal of Operational Research, Vol. 2, No. 6, pp. 429–444.

Clermont, M. (2016), Effectiveness and efficiency of research in Germany over time: An analysis of German business schools between 2001 and 2009, Scientometrics, Vol. 108, No. 3, pp. 1347–1381.

Clermont, M./Dirksen, A./Dyckhoff, H. (2015), Returns to scale of business administration research in Germany, Scientometrics, Vol. 103, No. 2, pp. 583–614.

Clermont, M./Dirksen, A./Scheidt, B./Tunger, D. (2017), Citation metrics as an additional indicator for evaluating research performance? An analysis of their correlations and validity, Business Research, Vol. 10, No. 2, pp. 249–279.

Costas, R./Zahedi, Z./Wouters, P. (2015), Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective, Journal of the Association for Information Science and Technology, Vol. 66, No. 10, pp. 2003–2019.

Dyckhoff, H. (2006), Produktionstheorie: Grundzüge industrieller Produktionswirtschaft, 5. Aufl., Springer.

Dyckhoff, H. (2018), Multi-criteria production theory: Foundation of non-financial and sustainability performance evaluation, Journal of Business Economics, Vol. 88, No. 7, pp. 851–882.

Dyson, R. G./Allen, R./Camanho, A. S./Podinovski, V. V./Sarrico, C. S./Shale, E. A. (2001), Pitfalls and protocols in DEA, European Journal of Operational Research, Vol. 132, No. 2, pp. 245–259.

Eisenführ, F./Weber, M./Langer, T. (2010), Rational Decision Making, Springer.

Emrouznejad, A./de Witte, K. (2010), COOPER framework: A unified process for non-parametric projects, European Journal of Operational Research, Vol. 207, No. 3, pp. 1573–1586.

Eysenbach, G. (2011), Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact, Journal of Medical Internet Research, Vol. 13, No. 4, Art. e123.

Garfield, E. (1996), When to cite. The Library Quarterly, Vol. 66, No. 4, pp. 449–458.

Haustein, S./Peters, I./Bar-Ilan, J./Priem, J./Shema, H./Terliesner, J. (2014), Coverage and adoption of altmetrics sources in the bibliometric community, Scientometrics, Vol. 101, No. 2, pp. 1145–1163.
Holmberg, K./Thelwall, M. (2014), Disciplinary differences in Twitter scholarly communication, Scientometrics, Vol. 101, No. 2, pp. 1027–1042.
In der Smitten, S./Jaeger, M. (2012), Ziel- und Leistungsvereinbarungen als Instrument der Hochschulfinanzierung, HIS Hochschul-Informations-System GmbH.
Jongbloed, B./De Boer, H./Kolster, R./Kottmann, A./Vossensteyn, H./Benneworth, P./Cremonini, L./Lemmens-Krug, K. (2015), Performance-based funding and performance agreements in fourteen higher education systems, Consortium for Higher Education and Research Performance Assessment.
Kousha, K./Thelwall, M. (2019), An automatic method to identify citations to journals in News stories: A case study of UK newspapers citing Web of Science journals, Journal of Data and Information Science, Vol. 4, No. 3, pp. 73–95.
Kromrey H./Roose J./Strübing J. (2016) Empirische Sozialforschung: Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung mit Annotationen aus qualitativ-interpretativer Perspektive, Vol. 13, Konstanz.
Melo, A. I./Figueiredo, H. (2019), Performance management and diversity in higher education: An introduction, Tertiary Education and Management, Vol. 26, No. 3, pp. 247–254.
Mittermaier, B. (2020), Peer review and bibliometrics, in: *Ball, R.* (Ed.), Handbook Bibliometrics, de Gruyter Saur, pp. 77–90.
Mohammadi, E./Thelwall, M./Kousha, K. (2016), Can Mendeley bookmarks reflect readership? A survey of user motivations, Journal of the Association for Information Science and Technology, Vol. 67, No. 5, pp. 1198–1209.
Ortega, J. L. (2016), To be or not to be on Twitter, and its relationship with the tweeting and citation of research papers, Scientometrics, Vol. 109, No. 2, pp. 1353–1364.
Priem, J./Taraborelli, P./Groth, C./Neylon, C. (2010), Altmetrics: A manifesto, in: http://altmetrics.org/manifesto, retrieved on 15.11.2021.
Rassenhövel, S. (2010), Performancemessung im Hochschulbereich: Theoretische Grundlagen und empirische Befunde, Gabler.
Rhaiem, M. (2017), Measurement and determinants of academic research efficiency: A systematic review of the evidence, Scientometrics, Vol. 110, No. 2, pp. 581–615.
Robinson-Garcia, N./Costas, R./Isett, K./Melkers, J./Hicks, D. (2017), The unbearable emptiness of tweeting: About journal articles, Plos One, Vol. 12, No. 8, Art. e0183551.
Thelwall, M. (2018), Early Mendeley readers correlate with later citation counts, Scientometrics, Vol. 115, No. 3, pp. 1231–1240.
Tunger, D./Clermont, M./Meier, A. (2018), Altmetrics: State of the art and a look into the future, in: *Jubi, M.* (Eds.), Scientometrics, Intechopen.
Tunger, D./Meier, A./Hartmann, D. (2017), Altmetrics Feasibility Study, in: http://juser.fz-juelich.de/record/851696, retrieved on 15.11.2021.
Van Noorden, R. (2014), Online collaboration: Scientists and the social network, Nature News, Vol. 512, No. 7513, pp. 126.
Van Vught, F./Ziegele, F. (2010), U-Multirank: Design and testing the feasibility of a multidimensional global university ranking, Consortium for Higher Education and Research Performance Assessment.
Zahedi, Z./Costas, R./Wouters, P. (2017), Mendeley Readership as a filtering tool to identify highly cited publications, Journal of the Association for Information Science and Technology, Vol. 68, No. 10, pp. 2511–2521.

Altmetrics als zusätzliche Indikatoren zur Analyse der Forschungseffizienz von Universitäten? Eine konzeptionelle und empirische Studie

Es existiert eine Vielzahl empirischer Studien zur Messung der Forschungseffizienz von Universitäten. Dabei werden als Maßgrößen hauptsächlich bibliometrische Indikatoren verwendet, mit denen die Verbreitung und Wirkung der Forschung in der Fachcommunity quantifiziert werden soll. Weitere Ziele des Hochschulmanagements im Zusammenhang mit der Forschungstätigkeit, die sich auf andere Stakeholder wie die interessierte Öffentlichkeit beziehen, werden in diesen Studien in der Regel nicht adressiert. Vor diesem Hintergrund stellt sich die Frage, ob Altmetrics geeignet sind, solche vernachlässigten Ziele zu quantifizieren, um sie als zusätzliche Indikatoren in mehrdimensionalen Effizienzanalysen einzusetzen. Im vorliegenden Beitrag wird diese Frage auf Grundlage eines entscheidungstheoretischen Rahmens konzeptionell untersucht, indem das Potenzial von Altmetrics zur Quantifizierung von Forschungszielen diskutiert wird. Dabei stellt sich heraus, dass nur wenige Altmetrics die grundlegenden Anforderungen an Maßgrößen erfüllen. Die verbliebenen Altmetrics werden sodann für eine empirische Analyse unter Zugrundelegung der Data Envelopment Analysis genutzt, um den Einfluss auf die Forschungseffizienz von drei verschiedenen Fachbereichen zu untersuchen. Es zeigt sich, dass die Berücksichtigung von Altmetrics ceteris paribus zu keinen signifikanten Veränderungen in den Rangplätzen der betrachteten Universitäten führt.

JEL-Kennziffern: C14, C44, E23, I23