12-2021

A Discipline in Search of a Voice: A Corpus Linguistic Study of Evaluation Scholarly Literature

Aaron Wilson Kates
Western Michigan University, awkates@gmail.com

A DISCIPLINE IN SEARCH OF A VOICE: A CORPUS LINGUISTIC STUDY OF EVALUATION SCHOLARLY LITERATURE

by

Aaron Wilson Kates

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Interdisciplinary Ph.D. in Evaluation
Western Michigan University
December 2021

Doctoral Committee:

Michael Harnar, Ph.D., Chair
Tarek Azzam, Ph.D.
Daniela Schroeter, Ph.D.

A DISCIPLINE IN SEARCH OF A VOICE: A CORPUS LINGUISTIC STUDY OF SCHOLARLY EVALUATION LITERATURE

Aaron Wilson Kates, Ph.D.

Western Michigan University, 2021

Linguists use the concept of discourse community (DC) to describe the speech of groups of people associated by some extrinsic purpose or interest (Swales, 1990). One defining feature of a DC is a shared lexicon. Scriven (1994), in his discussion of evaluation as a discipline, addresses the need for evaluation to have a "core subject" which would then allow, among other things, the development of "concepts and language to deal with core problems," while noting that there exists a definite lack of clarity of language among evaluators (p. 147). Christie and Rose (2003) suggest that the lack of clarification around definitions among evaluators is not simply a concern for linguists, raising fears that without this clarity, there could arise "a range of folk theories of evaluation that drive considerable real-world practice" (p. 42). They call for "a focused study of the language of evaluation [that] might help us to answer these lines of inquiry" (p. 42). To date, such a "focused study" of evaluation language has not occurred.

The question as to whether the evaluation community in general and the community of evaluation scholars more specifically, constitutes a DC has never before been considered and it has implications for evaluation's standing as a field. The purpose of this study is to examine the issue of whether the evaluation community indeed possesses its own lexicon to the extent that would lend credence to its status as a DC.

This study explores the lexicon and linguistic patterns of evaluation scholars using a set of established methods from corpus linguistics. CEJA2019, an approximately four-million-word corpus of scholarly evaluation literature, is analyzed using the Sketch Engine corpus linguistic platform. A word list and keyword list are generated. Other exploratory linguistic analyses are conducted.

The results of this study indicate a low level of specialized vocabulary. What specialized vocabulary does exist mainly relates to taxonomical distinctions between different theories, models and approaches of evaluations. The findings also suggest more precision of use among evaluation scholars surrounding central evaluation vocabulary including "evaluation" and "assessment." This study lends credence to Shadish's (1998) assertion that "evaluation theory is who we are" and may allay the concerns of those who fear a lack of precision in evaluation definitions.

ACKNOWLEDGEMENTS

This dissertation is dedicated to the memory of Dr. Chris Coryn, my friend and mentor. All I have to say is, “thank you.”

To my committee members, Dr. Michael Harnar and Dr. Tarek Azzam, and Dr. Daniela Schroeter, thank you for your thoughtful suggestions, insights, and support through the writing process. Thank you for your swift actions that allowed me to finish as planned.

To Zach Tilton, if your friendship was the only thing to come out of this program, it would still be worth it.

To my mother and father, Thomas and Jennifer Kates, I thank you both for sowing the seed of curiosity. This would not have happened without it.

To my siblings, Mike and Shamita, I thank you for many years of love, laughter and conversation. May there be many more.

To my father-in-law and mother-in-law, Bill and Kathy Cobern, I thank you for the love and generosity toward me, Rebecca, Simone, and Nels. You made this possible.

To my wife, Rebecca, I thank you for your enthusiasm, encouragement and sacrifice in this process. You have walked with me every step of the way and should get an honorary PhD in evaluation for it.

To Simone and Nels, I thank you for being a source of light and hope in the world. Yes, I’m going to be a doctor now. Just not that kind.

Lastly, to Drs. Albert and Florence Hazzard, thank you for your legacy, which I and all other future researchers in your lineage will carry on.

Aaron Wilson Kates

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Background of the Problem

Whereas a dialect describes a way of speaking that is bound by a geographic and/or ethnic boundary, the concept of *discourse community* (DC) is used by linguists to describe the speech of groups of people associated by some extrinsic purpose or interest[1] (Swales, 1990). In describing a discourse community around purpose or interest, one can examine the linguistic patterns of groups of people separated geographically and ethnically from one another but who nonetheless develop some identifiable common patterns of *discourse*[2].

Swales established eight criteria for a group to be considered a discourse community (Swales 1990, 2016). A discourse community…

1. has a potentially discoverable set of common goals
2. has mechanisms of intercommunication
3. has participatory mechanisms used primarily to provide information and feedback

---

[1] The exact definition in Swales (1990): "Sociorhetorical networks that form in order to work towards sets of common goals" (p. 9).

[2] Discourse is identified as "language above and beyond the sentence" (Schiffrin, 2011, para. 1). This would include such topics as the communicative purposes, subject matter, and means of communication found in language (generally or specifically)

4. possesses one or more *genres*[3] in the communicative furtherance of its aims
5. has acquired (or is developing) some specific *lexis*[4] (abbreviations, jargon, proprietary terminology, etc.)
6. has a hierarchical structure, explicit or not, that regulates entry and advancement in the community. He also includes in this criterion the requirement that there be a sufficient number of members with a suitable degree of relevant content and discoursal expertise (presumably without which the structure and hierarchy could not exist).
7. has '*silential relations*'[5]: "a sense of things that do not need to be said or spelt [sic] out in detail in either words or writing" (Swales, 2016, p. 9).
8. "develops horizons of expectation, defined rhythms of activity, a sense of history, and value systems for what is good and less good work" (p. 9).

These criteria have been extensively debated since their establishment. For instance, Borg (2003) questions the need for a discourse community to possess a common set of goals, objecting that "a 'public discourse community' cannot have shared goals, and more crucially, a generalized 'academic discourse community' may not have any shared goals or genres in any meaningful sense" (p. 399). He also questions whether the discourse of "such nebulous communities can be described in meaningful terms" (p. 399).

---

[3] Swales (1990) defines genres as "classes of communicative events which typically possess features of stability, name-recognition, and so on" (p. 9).

[4] This term in Swales (1990) and other linguists is essentially parallel to the concept of vocabulary, and often focuses on phrases, abbreviations, or jargon.

[5] Silential relations refers to means of unspoken communication in a discourse community. An example from the evaluation field could be use of a headline such as "RFP: Workforce Development Program." The underlying message is widely understood as, "We have an opportunity for an external evaluation concerning a workforce development program. If you are interested, please apply."

Swales himself has noted such difficulties with the concept of discourse community (Swales, 2016). However, its enduring potential in the literature speaks to its utility for describing the discourse of academic, professional, or other associations. While there continues to be some debate about the parameters of discourse communities, inquiry into these discoursal features "provides insight into what at first sight might seem standard, ordinary and predictable" (Swales, 2016, p. 10). In terms of the evaluation community, specifically, one might ask why it is important to know about these language patterns that may seem self-evident. Scholars may, for instance, state that there is a lack of development of the evaluation vocabulary or that some terms lack standardized usage. Such conjectures need empirical testing. Otherwise, they are simply intuitions.

Scriven (1994), in his discussion of evaluation as a discipline, addresses the need for evaluation to have a "core subject," which would then allow, among other things, the development of "concepts and language to deal with core problems" (p. 147). Both of these characteristics are reflected in Swales' first and fifth criteria. Scriven is not the only one who stresses the importance of common language for evaluators. For example, Schwandt (2015) points to the lack of a common definition for evaluation even among the most notable theorists as evidence of the fractious state of the field.

Statement of Problem

The question as to whether the evaluation community in general and the community of evaluation scholars specifically constitute a DC has never before been considered, and it has implications for evaluation's standing as a field. It is apparent that the field, mainly mediated by organizations such as the American Evaluation Association (AEA) and through publications such as the American Journal of Evaluation, meets many of the criteria of a discourse community. While there is room for debate and further inquiry, seven out of the eight DC

criteria above appear easily met. However, there remains the critical issue of whether the evaluation community indeed possesses its own lexicon to the extent that would lend credence to its status as a DC.

Study Objectives and Questions Investigated

The objective of this study is to explore and describe the evaluation lexicon.

1. To what extent does the scholarly evaluation literature possess a unique lexicon as compared to other fields?
2. What are the most significant keywords of the evaluation lexicon?

Significance of the Study

The exploration of the evaluation field as a DC could shed additional light on the state of theoretical and methodological unity or disunity of the evaluation field, as well as its status as a discipline.

Never before has an empirically-based effort to identify the lexicon of evaluation scholarship been attempted. Beyond the novelty of the study, the real significance of the study emanates from the potential applications. First, the study is significant in that it has the potential to create an empirically-based lexicon empowering those new to the field of evaluation to achieve a greater level of understanding of evaluation literature. Second, this study has the potential to reveal important findings related to the field of evaluation's status as a discourse community, and in turn, its level of development as a stand-alone discipline. Specifically, generating an empirically-based evaluation word list could illuminate the extent to which the evaluation field is generating unique concepts and tools. Additionally, any lack of important evaluation-specific vocabulary in

the word list could give evidence of a lack of awareness of or engagement with these concepts and tools in the evaluation community – even among academics.

In relation to this last point, the results of this study are likely to be quite open for interpretation. For instance, what would a lack of a developed set of discipline-specific technical lexis indicate? One could argue that the discipline is in an immature stage of development or even that it does not exist. Alternatively, it could simply be evidence that a more specified set of technical terminology (such as that proposed in Scriven (1991)) has not had the chance to infuse the scholarly evaluation community completely. Conversely, a highly developed set of technical terminology could yield a piece of evidence toward the evaluation discipline's maturity and status as a discourse community.

As inadequate as this study may be to answer such a line of inquiry fully, it will serve as a necessary first step toward doing so.

CHAPTER II

REVIEW OF THE LITERATURE

Foundational Concepts

Scriven (2013) describes evaluation as a basic cognitive process—indeed, perhaps the most essential cognitive process. By his definition, the process of determining "merit, worth, or significance" is thought to underly absolutely everything that we do as humans, and indeed any creature with any sort of cognitive capability.

Unsurprisingly, then, evaluation has deep roots entangled with human origins and history. Scriven (2013, p. 15) gives examples of early humans fashioning tools using a primitive form of product evaluation to gradually improve design. Evaluations such as this, with no overt or systematic application of values, criteria, or standards to a given situation are considered informal and arguably underly every human action.

While some formal and systematic evaluation was practiced by the Egyptian and Chinese imperial courts (Scriven, 1991), formal evaluation as a recognizable and cohesive social project began in the 1930s with the educational evaluator, Ralph Tyler (Stufflebeam & Coryn, 2014). As evaluation picked up speed as sub-fields in education, management, and manufacturing, there began to be some recognition of evaluation as an umbrella discipline for these pursuits, which finally began to coalesce in the last quarter of the twentieth century (Stufflebeam & Coryn, 2014). Some of the markings of this transformation were the emergence of professional networks and training programs for evaluators as well as a body of scholarly literature for the development and dissemination of evaluation

theory. As this occurred, the concept of evaluation as a discipline became a dominant thread in the scholarly literature.

Evaluation as a Discipline

Smith (1979) called for the establishment of evaluation as a discipline, stating this would be achieved through publicly establishing a body of knowledge based on "actual experience with the method", "public trials and applications of the method", and "the use of the method in a variety of settings and contexts" (p. 6). Here, Smith foreshadows later work by Scriven (1994) defining a discipline as a "core subject" that "knits the fields together with an overview and a map of connections…develops concepts and language to deal with shared problems, addresses threats to validity, converts solutions from one field for use by others, and moves the frontiers of foundations research forward" (p. 147). Much more recently, Patton (2018) addressed this topic, stating "scientific disciplines are distinguished by the overarching questions they ask and the body of knowledge that is built through inquiry into those questions" (p. 187). Given the diversity of definitions of evaluation and conceptions of exactly what it is that evaluators do, the precise nature of this core subject area is not agreed upon in the field. Despite noble attempts (such as Christie, Lemire, and Inkelas' (2017) attempt to differentiate evaluation from improvement science), full account of the exact differences between evaluation and other similar disciplines continues to elude the evaluation community at large. Christie and Rose (2003) suggest that the lack of clarification around definitions among evaluators is not simply a concern for linguists, raising concerns that without this clarity "a range of folk theories of evaluation" may emerge, "driv[ing] considerable real-world practice" (p. 42). They call for "a focused study of the language of evaluation [that] might help us to answer these lines of inquiry" (p. 42). To date, such a "focused study" of evaluation language has not occurred.

What Scriven, Patton, Christie, and Rose are describing and prescribing above is a certain type of social project that must be undertaken by a set of people. While a discipline is spoken of as a "body of knowledge" (Patton, 2018, p. 187), this body of knowledge exists only within and as defined by a collection of minds. Therefore, to further understand evaluation as a discipline, the means by which ideas are exchanged and coalesce must be considered and scrutinized.

Evaluation as a Discourse Community

The Introduction covered the history of the concept of Discourse Community (DC) as developed by Swales (1990, 2016) and the debate around the criteria by the likes of Borg (2003) and others. No further treatment will be given to the concept here, other than to state that the concept has never appeared in the evaluation literature – or if it has, it has never risen to a level of prominence.

Nevertheless, as stated before, it may be a useful way to further the examination of evaluation as a discipline. Of particular importance to this study is the focus on criteria one and five, those that deal with the clarification of a discernable goal and the development of a specific set of lexis. Again, these are echoes of Scriven's (1994) concerns around the "core subject" of evaluation and the "concepts and language" to deal with them (p. 147). Such questions as identifying core concepts and language may be greatly aided by applying methods established in linguistics—specifically those of lexicography, terminology, and corpus linguistics.

Lexicography, Terminology, and Corpus Linguistics

Lexicography, the study of dictionaries, is an ancient pursuit. The first lexicography in the English language was born out of necessity with the expansion of the Roman Catholic Church to Western Europe during the Middle Ages. The operation of monasteries was a bilingual affair, as the lingua franca of the British Isles collided with the Latin used in the sacred and academic texts of the era (Jackson, 2002). It took another thousand years for lexicography to take the leap of producing monolingual dictionaries, as the presence of loan words (words borrowed from other languages) increased (Jackson, 2002). Finally, dictionaries became more recognizable to the modern person in the eighteenth century, providing significant coverage of natural language use (Jackson, 2002). With the publication in 1755 of Samuel Johnson's *Dictionary*, the first of its kind in the English language to be compiled with any systematic rigor based on a collection of empirical data from written sources, the complete transformation from simple word lists to the modern dictionary was complete (Jackson, 2002).

Distinct from lexicography, Pearson (1998c) describes the discipline of *terminology* as emerging at the beginning of the Twentieth Century. As fields became more specialized, they began to require an ever-increasing level of precision for the language they used. Sager (1990) defines *terms* as lexical "items which are characterized by special reference within a discipline." By this definition, terms are specific linkages of concept and word that hold a certain importance or weight in a specific context. Linguists define *terminology* as the process of "collecting, disseminating and standardizing terms" (Pearson, 1998c, p. 10). It can also refer to the collection of terms produced by such a process.

Establishing the specific terminology of a language for a specific purpose (LSP) has historically been conducted from a top-down approach by which experts in a particular discipline determine the appropriate set of terms to be described (Pearson, 1998b). Such a process produces *standardized terms* (Pearson, 1998b). More recently, terminologists have increasingly turned to a

bottom-up approach that focuses on identifying and defining terms both in language for general purposes (LGP) and LSP (Pearson, 1998b). This may be due mainly to the emergence of high-powered computers and the ability to index billions of words at the touch of a button. There are, however, significant theoretical reasons for doing so as well:

> 'Traditional' terminologists tend to study terms in isolation from text and to ignore context, even when the terms have originally been sourced in text while 'modern' terminologies pay attention to usage, albeit mainly in the context of term recognition and the retrieval of appropriate contextual fragments" (Pearson, 1998a, p. 2).

This raises the concern that the traditional view of terminology may not be grounded in an empirical understanding of how language is used "in the wild," and thus of reduced efficiency and utility.

This shift in focus and the new availability of high-powered computers has led to the establishment of *corpus linguistics* as one of the dominant discourse analysis methods within linguistics. Cheng (2011) defines corpus linguistics simply as "the compilation and analysis of corpora" (p. 6). Scott and Tribble's (2006) definition is a set of linguistic "methods using corpora of texts, whether written or spoken, that is to say genuine examples of language in use" (p. 3). They contrast this with the "established alternative [of] intuition—methods relying on the speaker's own knowledge of the language and what seems to sound like possible utterance compared to what jars or seems 'un-English'" (p. 3) (or in our particular case, 'un-evaluese'). Scott and Tribble (2006) make the case that these methods are not in and of themselves new—the emergence of new hardware, software, and databases allows the methods to be applied more liberally and to better effect.

A corpus linguistic study often produces a *word list*, a *concordance*, and a *keyword list*. Scott and Tribble (2006) define a word list as "a list of word-types" (p. 12). In the modern era, this list would be generated by a computer by documenting the total number of types and tokens appearing in a text or a set

of texts. In this context, a type would mean a particular form of a word, while a token would refer to the number of times a particular word appears.[6] Word lists are most usefully ordered by frequency to give the user a sense of the dominance of particular lexical types in a corpus or text (Scott & Tribble, 2006, p. 15). In a frequency-ordered word list, the topmost common 100-200 items tend to be "closed-set items, a weft of prepositions, determiners, pronouns, conjunctions, whose role is mostly to glue texts together by supplying grammatical information to a lexical warp of nouns, verbs, adjectives, and adverbs" (Scott & Tribble, 2006, p. 24). As one proceeds down the list, the words become more infrequent and specialized. In the middle tend to be "common nouns, verbs, adjectives" (Scott & Tribble, 2006, p. 25). At the bottom of the list, as one adds more data, is an ever-expanding collection of *hapax legomena*, types that only occur a single time in a given corpus (Scott & Tribble, 2006, p. 26). These tend to be foreign words, proper nouns, and typographical errors.

In generating a concordance, a computer program documents the frequency of any two or more types occurring in a text or corpus together. This allows the researcher to "identify patterns of unusually high co-occurrence. That is, where two word-forms are found together (co-occur) more often than chance would predict" (Scott & Tribble, 2006, p. 33). Among other things, these concordances can be used to identify groupings of words that may have mental linkages in language users (Scott & Tribble, 2006, p. 40). *Multi-word units* (MWUs), are repeated strings of types that occur most frequently (Scott & Tribble, 2006, p. 41). Identification of these items (also known as *n-grams* or *lexical bundles*) is useful for discovering terminology that is based in MWUs rather than in stand-alone types.

---

[6] The following illustration may elucidate the definitions of and distinctions between types and tokens. Take the sentence, "I am what I am." The sentence consists of three distinct word forms, or types: "I", "am", and "what". However, the sentence consists of five words in total, yielding a token count of five. In short, the type count is the number of unique words, while the token count is the raw tally of all words in the text, repeated or not.

The most useful product of a corpus analysis for this dissertation is identifying keywords in a corpus. Keywords are identified by Scott and Tribble (2006) as "items of unusual frequency in comparison with a reference corpus of some suitable kind" (p. 54). This greater-than-expected frequency of particular types is referred to as *keyness*, defined as "a quality words may have in a given text or set of texts, suggesting that they are important, they reflect what the text is really about, avoiding trivia and insignificant detail. What the text 'boils down to' is its keyness, once we have steamed off the verbiage, the adornment, the blah blah blah" (Scott & Tribble, 2006, pp. 55-56). The level of keyness for a particular type is calculated by tabulating the *type-token ratio* (TTR) in both the corpus of interest and a reference corpus (Gilmore & Millar, 2018). A log-likelihood or chi-squared test has traditionally been employed in such software platforms as WordSmith tools to determine the keyness value (Baker, Hardie, & McEnery, 2006). The log-likelihood keyness value is "a statistic which, similar to chi-square, compares observed and expected values for two data sets but does not make assumptions of normal distribution" (Chiba, Millar, & Budgell, 2010).

Another approach to calculating keyness is known as "simple maths for keywords", which dismisses the notion of using the log-likelihood or Fisher's exact Test in linguistics, "since all it serves to do is to disprove a null hypothesis – that language is random – which is patently untrue" (Kilgarriff, 2009, p. 2).

Simple maths for keywords operates off the following formula:

$$\frac{fpm_{rmfocus} + N}{fpm_{rmref} + N}$$

Where $fpm_{rmfocus}$ and fpm_{rmref} are the normalized frequency of words in the focus corpus and reference corpus, respectively. N is a "smoothing parameter," normally set to a default value of 1. A smaller value (e.g., .001) focuses on rarer

terms, while a larger value (e.g., 1,000) focuses on more common terms. This is the approach employed in the online corpus linguistic tool, Sketch Engine.

The main utility of a list of keywords is to give the user an idea of the 'aboutness' of a text (Baker et al., 2006; Gilmore & Millar, 2018; Jablonkai, 2010; Pearson, 1998a; Scott & Tribble, 2006). This 'aboutness' is not only helpful in determining the nature of a text or corpus; it is also useful for determining the intangible qualities of the culture that produced the language being studied. Stubbs (2010) states, "keywords are the tips of icebergs: pointers to complex lexical objects which represent the shared beliefs and values of a culture" (p. 23).

Selection of Corpora

Rea-Rizzo (2010) defines a linguistic corpus as "a collection of written and/or oral naturally occurring texts" which is purposefully selected according to some sort of criteria as a means of representing a particular type of language. (p. 3). She particularly emphasizes the necessity of purposeful and systematic curation of corpora (p. 2), as representation is central to the goals of corpus linguistics.

Corpus linguists have set forth a variety of principles and guidelines to follow when establishing a linguistic corpus. Sinclair (2005) established the following ten guidelines as presented in Rea-Rizzo (2010):

1. The contents of a corpus should be selected without regard for the language they contain, but according to their communicative function in the community in which they arise.
2. Corpus builders should strive to make their corpus as representative as possible of the language from which it is chosen.
3. Only those components of corpora which have been designed to be independently contrastive should be contrasted.
4. Criteria for determining the structure of a corpus should be small in number, clearly separate from each other, and efficient as a group in delineating a corpus that is representative of the language or variety under examination.

5. Any information about a text other than the alphanumeric string of its words and punctuation should be stored separately from the plain text and merged when required in applications.
6. Samples of language for a corpus should wherever possible consist of entire documents or transcriptions of complete speech events, or should get as close to this target as possible. This means that samples will differ substantially in size.
7. The design and composition of a corpus should be documented fully with information about the contents and arguments in justification of the decisions taken.
8. The corpus builder should retain, as target notions, representativeness and balance. While these are not precisely definable and attainable goals, they must be used to guide the design of a corpus and the selection of its components.
9. Any control of subject matter in a corpus should be imposed by the use of external, and not internal, criteria.
10. A corpus should aim for homogeneity in its components while maintaining adequate coverage, and rogue texts should be avoided.

Most of these guidelines pertain to the notion that a corpus is meant to be representative of a certain population of speech events (see numbers 1, 2, 4, 6, 8, 9, 10). The issue of representativeness in linguistic corpora is heavily emphasized in the work of many others (Biber, Connor, & Upton, 2007; Cheng, 2011; Pearson, 1998a; Rayson, Berridge, & Francis, 2004; Rea-Rizzo, 2010). Biber, Conor, and Upton (2007) note that the need for representativeness in corpus linguistics is "of course no different than any other quantitative research in the social sciences, where there is always a concern that the 'sample' being studied represents the larger target 'population'" (pp. 17-18). They identify two potential barriers to effectively collecting a representative corpus (p. 18):

1. Corpora are often designed for general use rather than a specific study. As a result, the population being represented can be relatively general, such as newspaper language, or even an entire language.

2. Researchers sometimes choose to use a corpus just because it is publicly available, with little consideration of whether that corpus actually represents the target population being investigated.

Because of these issues, bounding the population of language to be examined is of great importance to ensure unbiased and equal access to all speech events contained therein. Biber, Conor, and Upton (2007) state that one of the best ways to achieve this is to compile specialized corpora, which "represent a narrowly defined genre" (p. 18). Cheng (2011) distinguishes between general corpora, those which "aim to examine patterns of language use for a language as a whole, and specialized corpora [which] are compiled to describe language use in a specific variety, register, or genre" (p. 166). General corpora are normally quite large, with token counts numbering in the hundreds of millions, or even billions (Cheng, 2011). Specialized corpora are generally much smaller, ranging from thousands to tens of millions of tokens (Cheng, 2011). Some corpora, when seen to have multiple categories of speech, are divided into sub-corpora. For instance, Biber, Conor, and Upton (2007) describe a study of non-profit fundraising letters that is divided into sub-corpora based on the type of non-profit. Another is described as a corpus of biology scholarly literature that is arranged by the sub-disciplines found therein (e.g., evolutionary, cellular, micro) (Biber, Conor, & Upton, 2007). Others (Biber et al., 2007; Chiba et al., 2010) perform more targeted studies of academic writing without identifying specific sub-corpora.

Applications of Corpus Linguistics

The primary use of corpus linguistics, especially lexicography, has been the identification of essential vocabulary for language learners (Brezina & Gablasova, 2015, 2017; Chiba et al., 2010; Gilmore & Millar, 2018; Jablonkai, 2010; Rea-Rizzo, 2010; Scott & Tribble, 2006). These studies primarily focus on producing

empirically-based lists of words and phrases necessary for a language learner to achieve facility in a particular lexis. Disciplines investigated in this way include plumbing, nursing, medicine, law, pharmacy, and international relations (Bancroft-Billings, 2020; Coxhead & Demecheleer, 2018; Grabowski, 2015; Jablonkai, 2010; Nguyen Le & Miller, 2020; Staples, 2019). Others also use corpus analysis to provide insights as to the nature of a particular discipline. For instance, Gilmore and Millar (2018) used their sub-corpus data to calculate the relatedness of subdisciplines of civil engineering. Again, many emphasize the importance of keyword analysis in determining the "aboutness" of a particular corpus (Pearson, 1998b; Scott & Tribble, 2006). For instance, Chiba, Millar, and Budgell (2010) found that midwifery literature "focuses on the interaction of mother, child and care-giver, processes related to birth and the importance of holistic care and client—care-giver interactions. Anatomical and pathological terms are uncommon" (p. 81).

Using corpus-based lexicography, analyzing the different 'levels' of lexis in a body of speech events can be a useful tool for drawing inferences regarding a particular discipline. Jablonkai (2010) synthesizes the different types of lexis into three categories: *technical lexis*, "highly specialized lexical items with no semantic ambiguity"; *semi-technical lexis*, which is "general lexis with a higher frequency in specialized texts" but may also have "specific, restricted meaning in certain disciplines; *general lexis*, which is comprised of "general content words" and "function words" (p. 46). Using these categorizations, one could calculate the ratios belonging to each category and draw inferences regarding the specialization of the language.

Linguistic Inquiry in Evaluation

The concept of DC and language use in evaluation scholarship has not been studied to any empirical extent in the research on evaluation literature. However, there have been some significant nods to the general importance of language to

evaluators. For instance, Patton (2000) focuses on issues of ambiguity in the language evaluators use and the role that language has in shaping the reality and practice of evaluators. Patton raises the issue of evaluation jargon, asserting the importance of "examin[ing] how it is used in a specific context, by whom, to communicate what, with what purposes and effects" (p. 9). Madison (2000) reviews how language use among evaluators has sociopolitical implications. The author uses the example of the term "at-risk youth" as used by evaluators as a case study, demonstrating how a shift in language in a particular political context led to the promotion of healthy youth development, rather than a pathology or problem-based model. Cabatoff (2000) conducted a case study to explore the implications of evaluators' language on their ability to influence policy change. Kaminsky (2000), speaking from his own experience working as an evaluation practitioner, examines how evaluators' use of metaphors influences how they view and practice their work. Shanker (2019) used discourse analysis to examine how language use in evaluation literature contributed to conceptions of race in the field and points to the importance of language and its effects on a community's thoughts, practices, and equity. However, the study focuses on one sole issue, that of race, and does not yield broader quantitative findings regarding the general lexicon of evaluation scholars.

Scriven (2013) addresses the importance of lexicography in determining the identity of a discipline:

> We know how to identify the lexical and behavioral footprints of most of these transdisciplines when we look for their practical applications in their own or other disciplines. In the case of evaluation, its lexical footprints are the use of terms like good and bad, right and wrong, better and best, etc., and the multitude of slightly less general evaluative terms… There are also scores of area-dependent derivatives, such as sick/well, intelligent/stupid, elegant/dowdy, ethical/unethical, valid/invalid. What may make this search – and any attendant conceptualization – harder in the case of evaluation is that, unlike most other transdisciplines, evaluation's vocabulary is not new and technical, hence not as

> easy to recognize as in the case of statistics, for example. There terms come from millennia of use in the common language.

Scriven's work in the lexicon of evaluation has mainly been constructive rather than descriptive. In his *Evaluation Thesaurus* (Scriven, 1991), he attempts to create a "large annotated glossary" (p. viii) relevant to what he terms as the "transdiscipline" of evaluation. Christie (2013) states that "Scriven has translated, developed, and shaped a vocabulary to describe some of the most fundamental aspects of our work—most famously, terms such as formative and summative evaluation" (p. 94). Scriven's work, according to Christie, has included "creat[ing] new terminology while defining and promoting existing concepts in relationship to evaluation" (p. 96). She asserts that many of these terms developed and endorsed by Scriven "are now normalized and central to the field" (p. 96).

Many governmental organizations such as the US Department of Energy and the US Centers for Disease Control and Prevention, and multilateral organizations, like the Organization for Economic Co-operation and Development have also recognized a pedagogical need for clarification of terminology of evaluation and have developed glossaries to this end (CDC, 2011; Education Endowment Foundation, n.d.; National Endowment for Financial Education, n.d.; OECD, 2002; Sportanddev.org, n.d.; Switchboard, n.d.; US Department of Energy, n.d.; US Fish and Wildlife Service, n.d.). All of these glossaries differ in size, scope, and terms included. This is the expected result when a vocabulary list is generated top-down by experts in the field who have differing views of the "aboutness" of evaluation. Dahler-Larsen and colleagues (2017) addressed issues of lexicon building as it relates to issues of translation and transferability of evaluation-related concepts to different linguistic contexts. This points to the reality that evaluation, a globalized phenomenon, is certainly not one discreet discourse community. Conversely, it is likely a complex set of nested discourse communities with varied levels of isolation and interaction, and each with differing cultural and linguistic backgrounds and contexts. This issue speaks to

both the importance of and the complicated nature of establish some agreed-upon lexicon. It also necessitates the acknowledgement of these complexities in any attempt to describe the lexicon at an aggregate level.

Aside from constructive efforts of lexicon building, there have been several attempts in the descriptive inquiry of normative language use in evaluation. Brown (2000) studied the linguistic dynamics of mixed-gender focus groups. Quantitative indicators such as talking time, topic raising, and leadership emergence were examined to reveal insights regarding gender-based power differentials in the focus groups.

Christie and Rose (2003) conducted a textual analysis of evaluation practitioners' writing. Their findings suggest a fundamental lack of clarity around evaluation terminology and definitions, leading to their call for greater, linguistically-focused inquiry as stated above. This study seeks to be the first attempt to answer that call.

CHAPTER III

METHODS

Research Questions

This study seeks to explore issues around the evaluation lexicon. It is concerned with answering the following two research questions:

1. To what extent does the scholarly evaluation literature possess a unique lexicon as compared to other fields?
2. What are the most significant keywords of the evaluation lexicon?

The following subsection will detail the methods employed to appropriately answer each question

Research Design

The research questions upon which this study rests lead directly to a well-established set of methods for identifying keywords and describing a lexicon. The first set of operations, including the development of the keyword and *n*-gram lists and the coverage analyses, aim to describe the lexicon and keywords in the evaluation literature. These methods include building a focused corpus, selecting reference corpora; generating word, keyword, and *n*-gram lists; and conducting coverage analyses. A sensitivity analysis is also described, which was used to assess the reliability of the results.

Sketch Engine, the main software package used for this study, includes a suite of exploratory language tools. As this study is the first of its kind in evaluation scholarship, the opportunity was taken to implement a small set of these to understand further the language used in evaluation scholarship. These methods include word sketches, word sketch difference, and thesaurus tools. The keyword, *n*-gram and coverage procedures are fully sufficient for answering the research questions. However, these exploratory language tools will contribute to a richer and more nuanced understanding of research question 1.

Building a Focus Corpus

The most critical aspect of a corpus linguistic study is that of constructing a quality focus corpus. This corpus will serve as the bedrock of the study, enabling the exploration of both research questions. Biber, Conor, and Upton (2007) emphasize that just as in other quantitative research, the representative nature of a corpus is of high importance, as it affects one's ability to infer to a larger collection of literature, speech, or population of speakers or writers.

A selection of journals was identified due to their emphasis on program evaluation, in general, and their use in previous investigations and systematic reviews of evaluation practice (Coryn, Noakes, Westine, & Schroeter, 2011; Coryn et al., 2017), in particular. Moreover, placing boundaries around subject- or discipline-specific journals (e.g., education, public health, medicine) would be virtually impossible given the ubiquitous, transdisciplinary nature of evaluation (Coryn & Hattie, 2006; Scriven, 1994).

As Scott and Tribble (2006) recommend, this corpus is narrowly focused so as to limit the introduction of systematic bias. The alternative of building a corpus of scholars across disciplines who write about evaluation, or a corpus that includes more genres of evaluation literature were possible. However, this would have been much more complex to build, as one could imagine many forms of bias that could be introduced in the selection process. For instance, one might

search for articles that list evaluation as a keyword, potentially leaving out many others that address evaluation but simply call it something else. This could introduce significant errors since the nature of the speech events themselves would simultaneously be of interest and affect their very inclusion in that sample. Similarly, efforts to add in other evaluation genres, such as conference preceedings or blogs, would add challenges regarding achieving a balance that accurately reflects evaluation speech in general.

The current study is focused on speech events belonging specifically to evaluation scholars, who are identified as those individuals who publish in evaluation-focused journals. The 14 scholarly evaluation-focused journals identified for this study are:

1. *African Evaluation Journal (AEJ)*
2. *American Journal of Evaluation (AJE)*
3. *Canadian Journal of Program Evaluation (CJPE)*
4. *Educational Evaluation and Policy Analysis (EEPA)*
5. *Evaluation: The International Journal of Theory, Research and Practice (EIJTRP)*
6. *Evaluation and the Health Professions (EHP)*
7. *Evaluation and Program Planning (EPP)*
8. *Evaluation Journal of Australasia (EJA)*
9. *Evaluation Review (ER)*
10. *Journal of MultiDisciplinary Evaluation (JMDE)*
11. *New Directions for Evaluation (NDE)*
12. *Practical Assessment, Research and Evaluation (PARE)*
13. *Research Evaluation (RE)*
14. *Studies in Educational Evaluation (SEE)*

The second issue of bounding relates to the temporally volatile nature of speech. The interest of the current study is to describe the language of

evaluation scholarship in its current state. Questions related to shifts in speech patterns in evaluation scholarship, while relevant, are outside the scope of this study. To this end, the temporal focus of this study was bound to 2019, the year before this study was initiated. The population of scholarly evaluation-focused journals is relatively small compared to many other fields, making a time-bound census of all scholarly evaluation journal articles feasible. To this end, all articles from the 14 evaluation-focused scholarly journals for the year 2019 were collected. The corpus will be henceforth referred to as the Corpus of Evaluation Journal Articles 2019 (CEJA2019). This title follows the tradition of other corpus linguistic studies, which tend to name the study corpus. The title enables tracing the origin of a collection of speech events more easily in the literature and the search for other studies that may have been conducted with the corpus. The title also suggests the possibility for more evaluation-related corpora to be collected under similar nomenclature, contributing to an even larger evaluation-specific corpus that possibly expands the temporal frame or the types of speech events included.

The third issue in corpus selection relates to issues of balance. There is a potential concern that there may be a particular type of speech event, author, or another such aspect that is over-represented or under-represented in a corpus. The current corpus, for example, could over-represent the number of articles related to medical evaluation relative to that of engineering or product evaluation. However, when considering the population of interest for this study, it is the population of speech events in this particular genre, namely evaluation scholarship. If medical evaluation happens to be over-represented, that is not a systematic bias introduced by the researcher. It instead tells us something about those speech events – that medical topics are over-represented in the literature relative to other types of writing. Attempting to create a balanced corpus according to what we think the community of evaluation scholars looks like or *ought* to look like, we get less, not more, signal regarding the state of the evaluation literature. This, again, would play to the hands of those seeking a

top-down, prescriptive model of evaluation lexicon building rather than the bottom-up process this study intends to approach.

Collecting the Focus Corpus

All articles for 2019 were downloaded as PDF documents from the 14 evaluation-focused journals. PDF is a type of document encoding that stores the documents as images rather than as encoded characters. The articles were then screened, and any that were not deemed to be related explicitly to scholarly evaluation literature were discarded. These were limited to 11 editorials that that pertained to subscriptions, pagination details, or other such topics that were not evaluation focused, and therefore outside the scope of the study. What remained was a sample of 483 articles containing 4.02 million tokens[7] and 3.13 million words. Gilmore and Millar (2018), Pearson (1998b), and Rea Rizzo (2010) suggest that the recommended minimum sample size is one-million words. This corpus is over three times this threshold, making it more than sufficient for this study. Further information regarding the composition of the focus corpus is provided in Chapter IV.

Preparing the Focus Corpus

After selection and collection of the focus corpus it was necessary to prepare the data for analysis. The online utility PDF2go.com was used to convert all 483 PDF articles into TXT files suitable for uploading to Sketch Engine. Corpus linguistic software packages traditionally use TXT files for compiling word lists and conducting other analyses, as plain text files are smaller allowing for more agility and speed during processing. PDF2go.com uses optical character recognition

[7] See footnote on pg. 10 for a distinction between types, tokens, and words.

(OCR) to convert pdf to plain text. This process resulted in some errors. Many of these errors were common and occurred throughout the corpus. To correct these, WordSmith Tools, a corpus linguistic software, was used to identify common typographical errors. Notepad++, a text editor which allows the user to manipulate code and perform batch searches and edits across a large collection of text documents, was used to find and replace these dominant errors. This was done in an iterative process until typographical errors were no longer prominent on the word list. Notepad++ was also used to batch-remove soft-hyphens, mid-sentence paragraph breaks, and any other errata left from converting PDFs to plain text. Reference sections were also removed from the articles to limit junk text. Upon close examination, some errors still carried over, but to a minimal extent.

Notepad++ was also used to eliminate data not directly related to the content. This includes reference sections, abstracts, and, where possible, headers and tables.

Once this process was complete, all .txt files were uploaded to Sketch Engine for compilation. Each journal was uploaded as a subcorpus for additional data on the size and description thereof.

Selecting Reference Corpus

As the identification of keywords and key *n*-grams require the identification of those lexical items that appear more than expected in a corpus (Baker et al., 2006), there needs to be a body of language to which the focus corpus can be compared. This body of language is called the reference corpus. Additionally, the use of a similar reference corpus in all respects except the subject matter is, according to Scott (2021), of high importance through the keyword and key *n*-gram procedures. “If your reference corpus (RC) is very similar to your study corpus (SC) in lexis (though bigger, maybe better stratified and presumably more representative), the KW (key word) procedure should throw up specifics of the

topics covered in SC. If RC is scholarly lit but not mainly from the SC's field, you should get lexis of that field too" (personal correspondence with Scott, February 9, 2021).

Finding a suitable reference corpus was a significant challenge for this study, as there are few large general academic corpora available. The Coxhead (2000) Corpus is large, but not accessible for researchers. The Directory of Open Access Journals (DOAJ) corpus, which includes 2.7 billion words and 3.3 billion tokens is available on Sketch Engine. However, it is, as its name suggests, open access – a feature unlike CEJA2019 and therefore disqualifying. The Corpus of Contemperary American English (COCA) and the British National Corpus (BNC) are both general language corpora, rendering them not ideal for the keyword and key *n*-gram procedures. The Michigan Corpus of Upper-Level Student Papers (MICUSP) is comprised of student papers, not scholarly journal articles.

The most suitable corpus of international English-language academic journal articles is the Corpus of Academic Journal Articles (Kosem, 2010). CAJA is international and is sufficiently large. The corpus contains 13,115 documents with 94.4 million tokens, 3.3 million sentences, and 1.4 million unique words. This corpus is available to academic researchers by request for analysis on Sketch Engine. The raw data cannot be downloaded but is available for analysis in the application, necessitating using Sketch Engine for this study rather than WordSmith.

The keyword and key *n*-gram procedures are run on the Sketch Engine platform. Therefore, for these operations, the raw data does not need to be downloaded. However, some of the coverage analyses described below require an exhaustive word list, so another sufficiently similar corpus will need to be substituted for those portions[8].

[8] Specifically, the coverage of the CAJA corpus by the word lists produced from the keywords procedure (the Evaluation Word List and the Evaluation Specific Word List) can be calculated. This is because the table of the keywords from CEJA2019 includes the word counts of all keywords from the CAJA corpus. Word lists not generated from CEJA2019 do not have the

Analytic Methods

The previous section detailed the rationale and procedure behind selecting and collecting linguistic data for this study. The following sections detail the methods for analysis. Data analysis was conducted exclusively in Sketch Engine and Microsoft Excel. As this is the first study of its kind in the evaluation field, any findings will serve as enlightenment on current issues in the field and a springboard for continuing corpus linguistic research.

Word List

The word list procedure generates an exhaustive list of lemmas[9] present in the corpus. In contrast to the keyword list as described in the next subsection, the lemma list is not generated based on any measure of keyness. It is instead derived from the program counting each appearance of every single lemma and arranging them from most common to least common. Lemmas occurring as few as one time occur on the list. This lemma list is used in two ways: First, the raw count of lemmas on the list give a sense as to the size of the lexicon of the corpus (see results section). Second, the lemma list is combined with the New Academic Word List and the New General Service List to conduct a lexical diversity analysis

---

associated word counts from the CAJA corpus, meaning that another corpus that does allow for download of an exhaustive word list, such as MICUSP, needs to be used.

[9] *Lemmas* are word groupings or families, meaning different forms of the same word. For instance, "are" and "is" would all be counted as belonging under the lemma, or parent verb, "be." Lemmas are used instead of words in this analysis because the coverage analyese described in following sections employ other word lists, namely the New Academic Word List and the New General Service List. These also use lemmas. This means that this study also needs to use lemmas, so as to maintain consistency.

(see pg. 30). Both of these pieces of information are used to answer research question 1.

The following settings in Sketch Engine produced a lemma list that excluded nonwords and was non-case-sensitive: “find: lemmas,” “all,” and “A=a,” “result format: simple list.” All other options were left as the default.

Keyword List

Whereas the word list procedure produced an exhaustive list of lemmas arranged by raw frequency, the keyword list is limited in size and designed to point toward both the subject matter and the uniqueness of the focus corpus. In corpus linguistics, keywords “describe lexical items that occur more frequently in the target corpus than would be expected by chance,” when compared to a larger reference corpus (Baker et al., 2006, p. 97), thus “represent[ing] what makes a corpus unique or different, and they can often provide a clear indication of what a set of texts is about” (Gilmore & Millar, 2018, p. 5). Hence, the corpus linguistics approach of keyword identification is useful for researchers wishing to know what distinguishes the field of evaluation from general scholarly literature.

This procedure required the use of both Sketch Engine and Microsoft Excel. In the Sketch Engine settings, a smoothing factor[10] of $N = 1000$ was used to produce a list of keywords that was balanced to focus on more typical words. A minimum frequency of five was required as a threshold to weed out terms of less common usage. Maximum items were set to 1,300 to allow for deletion of typos and errata after the fact. Options “at least one alphanumeric” and “only alphanumeric” were selected to ensure that words were only comprised of letters

[10] As stated in the literature review, N is a mathematical symbol for a smoothing factor used in the simple maths keyword calculations. It can be confused with a sample size, but it is not related in this case.

and numbers and not other symbols. The procedure was executed and the resulting table was exported into Microsoft Excel.

A dispersion metric that places a threshold on the minimum number of documents a word needs to appear in to be included on a keyword list are often applied in corpus linguistics. This ensures that a word does not appear to be more important than it is if it only appears in a small number of documents (Gilmore & Millar, 2018). For this study, a keyword had to appear in at least five documents in the focus corpus documents to remain on the list. Initially, this was set at five percent of the focus corpus, or 24 documents. However upon examining the output many obviously evaluation-specific items like "logframe" were excluded. Experimentation with the settings led to the current optimization. This filtering was done manually in Excel after exporting from Sketch Engine.

Obvious typographical errors and other errata such as "HTTP" were manually removed from the word list. Once these errors and words not meeting the minimum dispersion metric threshold were identified, they were manually removed in Excel. The top 1,000 words remaining are reported on the word list.

The top thirty words on Evaluation Word List and Evaluation Specific Word List were classified by the researcher according to Jablonkai's (2010) lexical levels. Ratios were calculated to determine the discipline-specificity of the word lists.

Word List Coverage Analyses

Calculating the coverage of a keyword list of various corpora is a method for determining the extent to which that list represents a corpus. Word lists can also be compared to determine the extent to which they overlap and are therefore similar or dissimilar. Below, the methods for conducting the overlap analysis for determining the Evaluation Specific Word List and calculating the coverage of the various word lists and corpora are described. These methods are directed

toward research question 1. They are specifically geared toward understanding the discipline-specificity of the evaluation lexicon.

Overlap Analyses. Jablonkai (2010) determined the discipline-specificity of the European Union Word List (EUWL) by determining the amount of overlap between the EUWL, the New General Service List , and the New Academic Word List. The extent to which the Evaluation Word List is evaluation-specific was determined by calculating the coverage of the New General Service List and New Academic Word List of the Evaluation Word List. This was done by counting the number of words in the Evaluaion Word List, and the New General Service List and New Academic Word List, respectively. Then, each coverage percentage was calculated by dividing the number of lemmas occurring in the New General Service List or the New Academic Word List by the number of words in the Evaluation Word List.

Words occurring in the Evaluation Word List but not the New General Service List nor the New Academic Word List were considered evaluation-specific and comprised the Evaluation Specific Word List.

Coverage Calculations. The coverage of a corpus by a lemma list was calculated using the following formula:

$$C = \frac{\sum k_i \ldots k_x}{w}$$

where C = coverage percentage, k = the keyword count, and w = the total number of words in the corpus. These calculations were performed using Microsoft Excel.

The procedure described above was used to determine the following coverages:

- CEJA2019 by the Evaluation Word List
- CEJA2019 by the Evaluation Specific Word List

- CEJA2019 by the New Academic Word List and New General Service List combined
- CEJA2019 by the Evaluation Specific Word List, New Academic Word List, and New General Service List combined
- CAJA by the Evaluation Word List
- CAJA by the Evaluation Specific Word List
- MICUSP by the Evaluation Specific Word List
- MICUSP by New Academic Word List and New General Service List combined
- MICUSP by Evaluation Specific Word List, New Academic Word List, and New General Service List combined

Because of lack of access to the entire CAJA word list in Sketch Engine, MICUSP was used as a proxy general scholarship reference corpus for the coverage analysis[11]. It is sufficiently large and represents scholarship from across disciplines at the University of Michigan. It lacks the size and international nature of CAJA. However, this particular coverage analysis simply acts as a test to demonstrate the difference in coverage in an evaluation-focused corpus versus a general academic corpus. Any inaccuracies here would be conspicuous and of less consequence to this study. Additionally, the coverage of Evaluation Specific Word List of CAJA is possible to calculate. This allows for comparison between CAJA and MICUSP to ensure that coverages may be assumed to be similar between the two.

Lexical Diversity Analysis. A lexical diversity analysis was conducted using the exhaustive lemma list described in the "word list" subsection (see pg.27). The

[11] As stated previously, since CAJA is proprietary and not available for download, only the word lists generated from CAJA can be used to calculate coverage of CAJA. This is because these word lists have the associated word frequency from CAJA. This is not true of other external word lists, including the New Academic Word List and New General Service List.

number of words that a reader must know to comprehend a certain percentage of words in a text has been treated by linguists as both a measure of readability and diversity (Coxhead, 2000). The threshold has traditionally been set at either 95% or 98% (Coxhead, 2000). Knowing the number of lemmas that are required to reach this level in the CEJA2019 Corpus, therefore, serves as an indicator of the diversity of the text, especially when compared to similar studies in other fields.

The steps for this analysis were carried out in Microsoft Excel using the exhaustive lemma list and the New General Service List and the New Academic Word List. First, lemmas occurring on the New Academic Word List and the New General Service List were removed from the exhaustive lemma list. Next coverage calculations using the same mathematics detailed in the previous subsection were employed to determine the percentave of CEJA2019 covered by New Academic Word List, New General Service List and the first 1,000, 2,000, 3,000 and 4,000 lemmas on the list.

The results of these calculations are used to draw conclusions toward research question 1 in that they indicate the size of the lexicon, as well as the diversity. In this case, a small number of words required to achieve the 95% readability threshold would indicate a small, highly specialized vocabulary. A large number of words required to achieve the threshold would indicate a large vocabulary with little specialization.

Key n-gram List

Similar to keywords, key *n*-grams, multi-word units that string together multiple words repeatedly and at a higher rate than expected compared to a reference corpus, can yield fiendings related to the “aboutness” of a literature (Scott &

Tribble 2006). These bundles themselves can also serve as evidence of a unique lexis. This process was executed to help answer the first research question regarding the uniqueness of the evaluation scholarship lexicon.

In this process, three- to four-word *n*-grams were generated to identify key co-occurring words in Sketch Engine. The "nest *n*-grams" option was selected, meaning that sub-*n*-grams are under the heading of the parent *n*-gram. For example, "call to action" and "a call to" would be sub-*n*-grams of "a call to action" and listed directly below the parent. The attribute of words (as opposed to lemmas or word families, e.g., the conjugations *is*, *be*, *were*, *are*) was selected. The same simple math described in the literature review was employed in this function. All other settings were the same as in the keyword operation.

Word Sketches

A word sketch is a representation of the grammatical and lexical associations of a word (Wang & Huang, 2017, p. 1). A grammatical association is the position in a sentence and the relationships a word may have. For example, a word sketch of the word "dog" may reveal that the top grammatical association is that the word acts as a direct object of a verb. Every such relationship, in order from most common to most uncommon, will be listed and will have an associated table.

Comprising each table are the lexical associations, which are, in order from most to least common, the words that most commonly appear in that particular lexical relationship. For instance, if the most common grammatical association for "dog" is as the direct object of a verb, the most common lexical association might be "walk", as in "walk the dog." This may be followed closely by "feed" and "pet", and so on.

Whereas keywords produce a shallow pool of knowledge about a broad set of data, word sketch is used to generate deep insights into using a single word. Word sketches were used to examine similarities and differences between the word "evaluation" in the CEJA and CAJA. This analysis contributes to

understanding of research question 1. In addition to the *n*-gram procedure this is a method that can yield additional insights as to the use of phrases which are part of the evaluation scholarship lexicon.

The process for conducting this operation is very simple. The appropriate corpus was selected for both, and then the search term "evaluation" was entered into the search box. No advanced options were selected.

Word Sketch Difference

The previous subsection detailed what a word sketch is. A word sketch difference allows for a word sketch comparison between two words, showing differences and similarities in use. This function was used to compare the usage of the words "evaluation" and "assessment". This method also contributes to understanding of question 1. Specifically, knowing the level of specification of use that evaluation scholars have for these pivotal words would add evidence for or against the existence of a specific evaluation language.

While a number of dimensions could have been explored, this study simply focused on one of the most common grammatical association: verbs for which "evaluation" and "assessment operate as objects. The results of this operation were downloaded as a visual directly from Sketch Engine.

While the mathematics and processing behind this operation is complex, the execution is simple. The following settings were used for word sketch difference: lemmas were compared, part of speech was a noun, and the minimum frequency was set to auto.

Thesaurus Tool

A thesaurus is a list of "synonyms or words belonging to the same category (semantic field)" ("Sketch Engine: Thesaurus," n.d.). Rather than simply identifying words with similar meaning, the Sketch Engine software creates a thesaurus where "the word sketches of all words with the same part of speech

are compared, and those that share the largest proportion of collocates are listed. The score given for each synonym indicates the percentage of shared collocates" ("Sketch Engine: Thesaurus," n.d.). While identifying words of similar meaning is useful as is done in traditional thesauruses, this method has the added benefit of yielding richer data by showing words that share similar relationships to other words as the search term. This is again a tool that can be used to illuminate the extent to which a vocabulary is specialized, especially when compared to other corpora. These data will contribute to learning associated with research question 1 in that any discernable differences will add to evidence of a disctinct lexicon in the evaluation scholarly literature.

This operation was executed for the word "evaluation" across several corpora and the results were compared. This function was performed by selecting the corpus to be used, followed by entering the desired search term. Results were compared between the CEJA, CAJA, and the British National Corpus (BNC) to detect any notable difference in usage between the evaluation literature, general academic literature, and the general English lexicon.

Sensitivity Analysis

All of the analyses detailed in the previous subsections rely on quality data for them to be trustworthy. One condition that could significantly affect the quality of this study, especially for coverage analyses, is the number of typographical errors in the data. A sensitivity analysis was performed to estimate the extent to which this may be an issue. To this end, one article from each journal was meticulously cleaned to remove or correct typographical errors. Many of these errors were introduced in the optical character recognition (OCR) phase of data processing. For instance, OCR often reads the letter "m" as "rn", leading to many occurrence of typos such as "rnake" or "rnove" versus "make" or "move". These fourteen articles were then uploaded to Sketch Engine as a corpus, and then the coverage analyses detailed above were performed. This process elucidates any

effects data quality may have on the coverage analyses and allows for calibration if any major effects are apparent.

Summary

The research questions upon which this study is based focus on the exploration of the lexicon of evaluation scholars. The processes detailed above serve to document and investigate the uniqueness of the evaluation lexicon and identify the dominant keywords in the evaluation lexicon. These aims are reflected in research questions 1 and 2, respectively.

The word list, keyword, and n-gram operations serve to identify the lexicon, while the coverage analyses, word sketches, word sketch differences and thesaurus operations explore the uniqueness of that lexicon. The sensitivity analysis serves as evidence as to the integrity of the focus corpus, the data upon which this study rests and from which it draws its credibility.

Limitations

This study, being the first of its kind to study the language of evaluation scholars, attempts to cover a lot of ground. As such, there are several areas in which limitations may be present.

First are challenges related to the sample. The data collected only represent a narrow time period. This limits the extent to which inferences regarding the language of evaluation scholars can be made in the past or present. Additionally, the genre selected is narrow. It is entirely possible that the lexicon of evaluation scholars may differ in conference proceedings, evaluation reports, or other genres.

Second, there were some technical limitations related to data transformation and quality. An unknown amount of error was added to the data in the process of OCR and conversion of articles from PDF to TXT.

Third, there are limitations in the methods themselves. For instance, the classification of keywords according to Jablonkai's (2010) levels of lexis was only conducted for the top thirty lemmas due to the limited time and resources for the study. The extent to which these findings regarding specialization can be inferred to the rest of the lemmas present on these lists is extremely limited. An alternate method would be to take representative samples of the lists and classify those. Limited time and resources made this infeasible for this study, but this weakness also represents an opportunity for the future.

Lastly, there are many different facets to the word sketches, word sketch differences, and thesaurus operations that were not explored. This means that there is plenty of room for additional pieces of data that could confirm or refute the findings of this study. Again, this represents both a challenge for this research as well as a future line of inquiry.

CHAPTER IV

RESULTS

The Corpus of Evaluation Journal Articles 2019 (CEJA2019)

The data collection process resulted in a respectably sized corpus deemed suitable for general corpus linguistics purposes related to evaluation scholarship. The CEJA2019 exceeds the minimum million-word standard set by many corpus linguists (Gilmore & Millar, 2018). The corpus includes 483 articles from 14 different evaluation-specific journals from 2019 (see Table 1). The corpus contains 4 million tokens, 3.1 million words, 114,848 sentences, and 151,692 unique words. The corpus is weighted toward European journals (53.83%) over American journals (38.10%) (See Tables 1 & 2). There are two journals from other regions, Oceana and Africa, comprising 5.59% and 2.48% of articles, respectively.

There is considerable variation in the size of each journal's contribution to the corpus. For instance, *Evaluation and Program Planning* includes 119 articles and makes up 25.14% of the corpus, compared to *Practical Assessment, Research, and Evaluation*, containing only nine articles, comprising 1.95% of the corpus.

Table 1

List of Included Journals by Region, Articles, and Tokens

Journal	Region	Articles	Tokens	Percent of Total Tokens Used
African Journal of Evaluation	Africa	12	81,238	2.02%
American Journal of Evaluation	North America	37	332,397	8.27%
Canadian Journal of Program Evaluation	North America	31	177,576	4.42%
Educational Evaluation and Policy Analysis	North America	23	328,154	8.16%
Evaluation	Europe	30	240,793	5.99%
Evaluation and the Health Professions	North America	24	173,368	4.31%
Evaluation Journal of Australasia	Oceana	27	942,42	2.34%
Evaluation and Program Planning	Europea	119	1,010,425	25.14%
Evaluation Research	North America	9	121,341	3.02%
Journal of MultiDisciplinary Evaluation	North America	14	83,842	2.09%
New Directions for Evaluation	North America	37	22,4082	5.57%
Practical Assessment, Research, and Evaluation	North America	9	78,271	1.95%
Research Evaluation	Eurpoe	38	33,6029	8.36%
Studies in Education Evaluation	Europe	73	73,7638	18.35%
Total		483	4,019,441	100.00%

Table 2

Corpus Composition by Region

Region	Articles	Article Percentage	Total Tokens Count	Token Percentage
North America	184	38.10%	1,519,031	37.79%
Europe	260	53.83%	2,324,885	57.84%
Oceana	27	5.59%	94,242	2.34%
Africa	12	2.48%	81,283	2.02%

Word List

Execution of the word list procedure produced a lemma list that includes 70,755 items. The table only contains two rows: “Item” and “Frequency”. The table is arranged by the “Frequency” column in order of greatest to least. At the top of the most prominent words are function words, starting with “the” with 179,972 appearances in the corpus. There are a few more specific lemmas near the top of the list, signalling their relative importance, such as “evaluation” (no. 11 with 24,440 occurrences) and “program” (no. 21 with 12,399 occurrences). At the bottom are lemmas only appearing once in the corpus, ending with “a-11”. Many of the lemmas at the bottom are typos, while others are simply infrequent words such as “urine”. The prohibitive size of this table makes its inclusion in the appendices unfeasable. The word list is available from the author upon request.

Lexical Diversity Analysis

The coverage of of CEJA2019 by the exhaustive lemma list items that did not appear on the New General Service List and New Academic Word List was calculated. Traditionally, linguists test the number of lemmas from the exhaustive lemma list combined with the New General Service List and the New Academic word list against the corpus from which the exhaustive lemma list was

created to determine the number of items required to cover 95% to 98% of the corpus (Coxhead, 2000). This is generally seen as a means of testing readability, but it also is an indicator of the lexical diversity of the corpus.

When combined with the New General Service List and the New Academic Word List, the top 1,000, 2,000, 3,000, and 4,000 lemmas yielded 91.70%, 93.42%, 94.44%, and 95.14% coverage, respectively. In contrast, Chiba, Millar, and Budgell (2010) crossed the 95% threshold with 2,000 additional items. Gilmore and Millar (2018) achieved a 92.4% coverage rate with only 650 off-list words, a level only approached (and not achieved) with the top thousand items from this list. Jablonkai (2010) found that their European Union Word List (EUWL) of 512 items combined with the first 2,000 word families of the General Service List achieved 93.5% coverage of their European Union corpus. These comparisons indicate a more extensive vocabulary is needed for a reader to comprehend the documents included in the CEJA2019 corpus than would be needed for other academic corpora. This suggests that the evaluation scholarship language has a larger lexicon, is more diverse, and is less discipline-specific than other academic languages.

Keywords

The keywords procedure produced a table of the top 1,000 keywords from CEJA2019 (see Appendix B). This is the Evaluation Word List. Table 3 includes details of the top 30 lemmas on the list. The table includes the item (lemma), frequency and relative frequency in the focus corpus (FC) and reference corpus (RC), keyness score, and lexical level classification (Jablonkai 2010). These are arranged by keyness score, the product of the simple maths calculation detailed in the literature review.

Table 3

The First 30 Lemmas on the Evaluation Word List

Item	Frequency (FC)	Frequency (RC)	Relative frequency (FC)	Relative frequency (RC)	Score	Lexical Level
evaluation	24,440	12,257	6,080.45	129.91	6.27	Semi-technical
program	12,399	18,967	3,084.76	201.02	3.40	Semi-technical
student	11,907	42,198	2,962.35	447.24	2.74	General
teacher	8,401	20,915	2,090.09	221.67	2.53	General
school	9,172	31,696	2,281.91	335.93	2.46	General
research	10,523	49,977	2,618.03	529.69	2.37	General
intervention	4,947	9,566	1,230.77	101.39	2.03	Semi-technical
outcome	5,378	18,155	1,338.00	192.42	1.96	Semi-technical
evaluator	3,884	470	966.30	4.98	1.96	Semi-technical
project	5,306	22,335	1,320.08	236.72	1.88	Semi-technical
assessment	4,466	12,880	1,111.10	136.51	1.86	Semi-technical
health	4,381	17,446	1,089.95	184.90	1.76	General
impact	4,648	21,317	1,156.38	225.93	1.76	Semi-technical
et	8,565	77,138	2,130.89	817.55	1.72	General
participant	5,170	31,088	1,286.25	329.49	1.72	Semi-technical
implementation	3,187	7,849	792.90	83.19	1.66	Semi-technical
community	4,646	28,802	1,155.88	305.26	1.65	General
education	4,031	22,977	1,002.88	243.52	1.61	General
programme	2,705	6,522	672.98	69.12	1.57	Semi-technical
learn	3,968	25,435	987.20	269.57	1.57	General
stakeholder	2,406	2,194	598.59	23.25	1.56	Semi-technical
team	2,917	10,067	725.72	106.70	1.56	General
practice	4,495	36,568	1,118.31	387.57	1.53	General
approach	5,178	47,740	1,288.24	505.98	1.52	General
item	3,368	19,933	837.93	211.26	1.52	Semi-technical
study	11,041	140,050	2,746.90	1,484.33	1.51	Semi-technical
how	5,767	59,079	1,434.78	626.15	1.50	General
survey	2,882	14,293	717.02	151.49	1.49	Semi-technical
focus	4,145	34,286	1,031.24	363.38	1.49	General
university	2,761	12,532	686.91	132.82	1.49	General

A smaller list, the Evaluation Specific Word List, includes 412 words derived from the Evaluation Word List that do not appear on the New General Service List (Browne, Culligan, & Phillips, 2013), nor the New Academic Word List (Coxhead, 2000). Table 4 details the top 30 words from the Evaluation Specific Word List. Since the Evaluaiton Word List is purely keyness-based, derived from the words that appear more frequently than expected, this list can be assumed to represent the "aboutness" of evaluation, while the Evaluation Specific Word List can be assumed to represent more of what makes the language of evaluation scholarship unique. For instance, the Evaluation Word List includes the word *approach*, while Evaluation Specific Word List does not. This would indicate a focus on approach and methodology as it relates to evaluation scholarship. In other words, it appears more frequently than expected in the evaluation literature. However, it is also a word present on both the New General Service List and New Academic Word List, and therefore not on the Evaluation Specific Word List. A general reader is likely to know this word, and therefore it is not deemed an evaluation-specific word. However, upon examining the top thirty words from both lists, it is readily apparent that there are few technical words on either. The words on the Evaluation Specific Word List that may be considered to have a specialized meaning particular to the evaluation field are *monitoring*, *evaluative, rubric*, *stakeholder,* and *evaluator.*

The top 30 words of the Evaluation Word List are comprised of 15 (50%) general and 15 (50%) semi-technical terms. The Evaluation Specific Word List is slightly more specialized, with 15 (50%) of the top 30 being general, 13 (43.33%) being semi-technical, and 2 (6.66%) technical.

Keyword Coverage Analyses

The Evaluation Word List, the more extensive thousand-word keyword list, was tested for coverage against the CAJA and the CEJA2019. Since the Evaluation Word List was created from the CEJA2019, this may seem redundant.

Table 4

The First 30 Lemmas on the Evaluation Specific Word List

Item	Frequency (FC)	Frequency (RC)	Relative frequency (FC)	Relative Frequency (RC)	Score	Lexical Level
evaluator	3,884	470	966.30	4.98	1.96	Semi-technical
et	8,565	77,138	2,130.89	817.55	1.72	General
programme	2,705	6,522	672.98	69.12	1.57	Semi-technical
stakeholder	2,406	2,194	598.59	23.25	1.56	Semi-technical
datum	6,642	80,212	1,652.47	850.13	1.43	Semi-technical
learning	2,320	9,613	577.19	101.88	1.43	General
training	2,657	15,590	661.04	165.23	1.43	General
funding	1,837	2,477	457.03	26.25	1.42	General
finding	2,919	23,557	726.22	249.67	1.38	Semi-technical
planning	1,828	7,516	454.79	79.66	1.35	General
sustainability	1,293	1,290	321.69	13.67	1.30	Semi-technical
teaching	1,578	7,816	392.59	82.84	1.29	General
engagement	1,265	3,997	314.72	42.36	1.26	Semi-technical
al	1,195	2,770	297.31	29.36	1.26	General
their	12,105	211,018	3,011.61	2,236.49	1.24	General
collaboration	1,047	2,438	260.48	25.84	1.23	General
DOI	778	156	193.56	1.65	1.19	General
rubric	782	533	194.55	5.65	1.19	Technical
rater	728	738	181.12	7.82	1.17	Semi-technical
competency	729	1,244	181.37	13.18	1.17	Semi-technical
understanding	1,477	16,536	367.46	175.26	1.16	General
math	688	549	171.17	5.82	1.16	General
organizational	982	7,709	244.31	81.70	1.15	General
respondent	1,053	9,586	261.98	101.60	1.15	General
collaborative	639	1,764	158.98	18.70	1.14	Semi-technical
evaluative	612	1,209	152.26	12.81	1.14	Technical
causal	776	4,928	193.06	52.23	1.13	Semi-technical
monitoring	679	2,969	168.93	31.47	1.13	Semi-technical
accountability	621	1,770	154.50	18.76	1.13	Semi-technical
meeting	791	5,387	196.79	57.09	1.13	General

However, the extent to which the word list covers the corpus demonstrates that it both is an accurate representation of the keywords in the corpus, and allows for comparison to the lexicons found in other corpora.

The analysis revealed that the Evaluation Word List covers 40.6% of the words in CEJA2019 and 20.8% of the CAJA. This suggests that the Evaluation Word List is indeed more representative of evaluation scholarship than general scholarship.

Table 5

Corpus Coverage by the Evaluation Word List

Corpus	Words	Word List Occurrence	Coverage
CAJA	78,970,229	16,436,585	20.81%
CEJA2019	3,125,612	1,269,552	40.62%

An analysis of the coverage of the Evaluation Word List by the New Academic Word List and New General Service List was conducted to determine the lemmas that are off-list, or do not appear on the New Academic Word List nor New General Service List. This resulted in the Evaluation Specific Word List. These are the terms that can be more unique to the evaluation discipline. This list has a much smaller coverage of the corpora, with 1.67% of CAJA, 2.04% of MICUSP, and 5.44% of CEJA2019. This is much smaller than similar studies such as Jablonkai (2010, p. 193), which found that their word list of 512 European Union (EU) terms covered 17.52% of EU texts. Chiba, Millar, and Budgell (2010) identified just 242 core words in midwifery texts that provided 6.91% of their midwifery corpus.

Combined with the General Service List[12] and Academic Word List, their word list provided 92.46% coverage (p. 78). Gilmore and Millar (2018, p. 8)

[12] General Service List and Academic Word List are older versions of the New General Service List and New Academic Word List

developed a list of 650 civil engineering keywords that covered 92.4% of their civil engineering corpus when combined with the New General Service List and New Academic Word List. The results of the coverage analysis of the Evaluation Specific Word List in combination with the New General Service List and the New Academic Word List yielded only 87.30% coverage of the CEJA2019 (see Table 6). This reveals that Evaluation Specific Word List has a relatively low coverage of CEJA2019 when compared to other discipline-specific word lists' coverage of similar discipline-specific corpora.

Table 6

Corpus Coverage by Combinations of the Evaluation Specific Word List (ESWL), New General Service List (NGSL), and New Academic Word List (NAWL)

Corpus	Coverage by ESWL	Coverage by NGSL and NAWL	Coverage by NGSL, NAWL, and ESWL
CAJA	1.67%	--	--
CEJA2019	5.44%	81.86%	87.30%
MICUSP	2.04%	81.76%	83.80%

Sensitivity Analysis

The sensitivity analysis described in Chapter III was performed to estimate poor data quality's effect on the coverage calculations described above. The cleaning of fourteen articles, one from each corpus journal, and subsequent coverage calculation produced minor changes. The coverage of the fourteen articles by New General Service List , New Academic Word List, and Evaluation Specific Word LIst was 90.85%. This indicates an overall coverage improvement of 3.55% when the data is cleaned to a greater extent. Coverage of the fourteen articles by the Evaluaiton Specific Word List was 5.40%, which is remarkably like the coverage of the entire uncleaned corpus by the Evaluation Specific Word List (5.44%). While this indicates that data quality does have some effect on the

coverage by the general lists, it yielded no effect on the coverage of evaluation-specific words. This means that most typos corrected were not evaluation-specific and were mostly included in the New General Service List and New Academic Word List.

Key *n*-grams

The key *n*-gram analysis produced a list of 1,000 lexical bundles or word clusters that occurred more frequently in the focus corpus than in the reference corpus. Table 7 lists the top 30 most significant *n*-grams. In addition to the keyword analysis, these *n*-grams add to an understanding of the "aboutness" of the evaluation literature. Notably, several of the phrases included in the top 30 list are directly related to Scriven's (1994) notion of evaluation as determination of merit, worth, or significance (e.g., *the importance of, the quality of)*. Twelve on the list are simply titles of evaluation journals that repeatedly appear in the literature (e.g., *New Directions for Evaluation, Studies in Educational Evaluation*). Two items are familiar evaluation-specific terms (i.e., *theory of change; and monitoring and evaluation*).

Word Sketches

The Word Sketch tool in Sketch Engine elucidates the use of a word in a corpus. This tool produces a set of tables listing lexical functions of a word in order of dominance. In the case of the CEJA2019 corpus, the Word Sketch function identified twelve distinct uses of the word "evaluation":

1. "modifiers of 'evaluation'"
2. "nouns modified by 'evaluation'"
3. "verbs with 'evaluation' as object"

Table 7

The Top 30 Key *n*-grams in CEJA2019

Item	Frequency (FC)	Frequency (RC)	Relative frequency (FC)	Relative frequency (RC)	Score
and program planning	908	0	225.90	0.00	1.23
evaluation and program	908	0	225.90	0.00	1.23
evaluation and program planning	904	0	224.91	0.00	1.23
in educational evaluation	653	0	162.46	0.00	1.16
studies in educational	648	0	161.22	0.00	1.16
studies in educational evaluation	648	0	161.22	0.00	1.16
theory of change	643	14	159.97	0.15	1.16
of the program	594	489	147.78	5.18	1.14
of the evaluation	547	185	136.09	1.96	1.13
of the intervention	414	295	103.00	3.13	1.10
new directions for	388	20	96.53	0.21	1.10
directions for evaluation	385	0	95.78	0.00	1.10
new directions for evaluation	385	0	95.78	0.00	1.10
the importance of	727	7678	180.87	81.38	1.09
as well as	1659	28819	412.74	305.44	1.08
in the evaluation	308	259	76.63	2.75	1.07
the implementation of	364	1573	90.56	16.67	1.07
american journal of	273	22	67.92	0.23	1.07
implementation of the	317	1041	78.87	11.03	1.07
journal of evaluation	269	0	66.92	0.00	1.07
american journal of evaluation	269	0	66.92	0.00	1.07
monitoring and evaluation	258	43	64.19	0.46	1.06
the quality of	414	3549	103.00	37.61	1.06
of the research	277	1024	68.92	10.85	1.06
can not be	232	189	57.72	2.00	1.06
as part of	405	4072	100.76	43.16	1.06
of the project	264	929	65.68	9.85	1.06
the purpose of	388	3784	96.53	40.11	1.05
extent to which	376	3667	93.55	38.86	1.05
to participate in	282	1557	70.16	16.50	1.05

4. “verbs with ‘evaluation’ as subject”
5. “‘evaluation’ and/or”
6. “prepositional phrases”
7. “adjective predicates of ‘evaluation’”
8. “’evaluation’ is a…”
9. “evaluation’s…”
10. “pronominal possessors of ‘evaluation’”
11. “… is a ‘evaluation’”
12. “verbs with particle ‘out’ and ‘evaluation’ as object”

The top-ranked use of “evaluation” in CEJA2019 as documented by Sketch Engine illustrates cases in which another word modifies “evaluation.” The “modifiers of ‘evaluation’” table in the Word Sketch tool details all ways in which another word modifies the word “evaluation.” Table 8 includes the first 30 of these modifications.

Modifiers can be distinguished into evaluation objects (evaluands), descriptors, discipline-specific terms, and others. Most of the items on the list can be classified as evaluation taxonomy, which is concerned with differentiating between the various theories, models, and approaches of evaluation.

Evaluation objects include, for example, “portfolio,” “course,” and “program.” Examples of descriptors include “rigorous,” “realistic,” and “future.” Examples of discipline-specific taxonomical delineations are “realist,” “empowerment,” and “responsive.” The rest (others) are simply errata, artifacts of the academic context in which the linguistic features of interest are embedded (e.g., “et,” “al”). This illustrates that the primary way evaluation is discussed in the evaluation literature is in the context of evaluation theories, models and approaches.

The Word Sketch tool was also used to generate the modifiers of “evaluation” in the CAJA corpus to compare and contrast differences in usage

Table 8

Modifiers of “Evaluation” in CEJA2019

Word	Frequency	Score
educational	613	11.28
program	353	9.91
research	227	9.81
realist	166	9.48
impact	159	9.30
participatory	127	9.07
process	112	8.87
research	150	8.70
formative	94	8.67
internal	93	8.55
external	90	8.52
course	84	8.45
al	78	8.35
teacher	82	8.19
ongoing	59	7.97
developmental	51	7.79
summative	50	7.79
adversary	48	7.75
responsive	48	7.73
independent	49	7.66
policy	52	7.6
et	50	7.56
empowerment	42	7.54
outcome	45	7.50
portfolio	39	7.44
future	42	7.37
realistic	37	7.37
module	35	7.29
program	35	7.26
rigorous	34	7.22
realist	32	7.16

between the general and evaluation scholarly literature. The top 30 modifiers of "evaluation" from CAJA are listed in Table 9. Similar to the CEJA2019 corpus, many specify the thing being evaluated, the evaluand (e.g., "performance," "function," "candidate," "program/me," "brand," "outcome," "customer," "query," "reliability," and "assessor"); descriptors ("comparative," "critical," "overall," "positive," "lazy," "systematic," "favorable," "prospective," "in-clinic," "follow-up," "quantitative," "negative," "subjective," "clinical," and "retrospective," and "driving"). Only two uses would be recognizable as discipline-specific evaluation terms (i.e., "empowerment evaluation," "economic evaluation").

This simple comparison of word sketches reveals a simple difference between evaluation and general scholarly literature: When modifying the word "evaluation," the literature is much more likely to do so in a manner that attempts to make discipline-specific taxonomical delineations, . The general scholarly literature has picked up on some of these terms (e.g., "empowerment evaluation") but is less likely to use these discipline-specific terms when modifying "evaluation."

Word Sketch Differences

Word Sketch Differences use the same methods as the Word Sketch tool, but in contrast, compares the usage of different words in a corpus. Words are compared by their different collocations on each type of usage as it appears in the word sketch. The best way to view these data are by using the graphics generated by Sketch Engine.

For the further exploration of the use of "evaluation" in CEJA2019, the usage of "assessment" was compared. These terms are often interchangeably used in the literature and in common speech, making them ripe for comparison, potentially teasing out precise delineations in usage.

The dimension on which the two words were compared was the primary usage of "evaluation" in CEJA2019: Associated modifiers (see above). On the

Table 9

Top 30 Modifiers of “Evaluation” in CAJA

Word	Frequency	Score
empowerment	191	11.12
performance	228	9.39
function	77	8.23
retrospective	41	8.18
assessor	21	8.06
candidate	36	7.92
economic	156	7.7
programme	19	7.7
program	31	7.66
clinical	78	7.64
subjective	41	7.59
negative	124	7.57
quantitative	47	7.55
brand	25	7.48
follow-up	23	7.47
reliability	19	7.43
outcome	23	7.4
in-clinic	18	7.23
prospective	23	7.19
favorable	22	7.14
query	19	7.13
experimental	80	7.08
systematic	32	7.07
lazy	16	7.04
customer	29	6.99
positive	94	6.98
overall	63	6.98
critical	58	6.98
driving	11	6.98
comparative	30	6.97

right are the words collocated (used alongside) with “evaluation,” and on the left are words collocated with “assessment.” In the center are words that are collocated more or less equally with both words. Figure 1 portrays verbs with “evaluation” or “assessment” as an object. Notably, there are some words (“commission,” “standardize,” and “manage”) that are only associated with “evaluation.” However, none are only associated with “assessment.” This suggests that when operating as a subject of a verb, the word “evaluation” may cover most (if not all) instances where the word “assessment” could be used. However, there are instances where the converse is not true, and “assessment” cannot be used in place of “evaluation.”

Two things are noticeable compared to the same visual with that of the parallel figure (Figure 2) generated from the CAJA corpus. First, apart from some exceptions, most collocations appear more or less equally alongside “evaluation” and “assessment. Second, there are far fewer collocations. This suggests that there is greater specificity for the distinction of assessment from evaluation in evaluation scholarship. How both “evaluation” and “assessment” operate as objects of a verb are much more diverse in the evaluation literature.

Thesaurus

The Thesaurus tool in Sketch Engine generates a list of words that belong in the same “semantic field” (are similarly collocated) (“Sketch Engine: Thesaurus,” n.d.) in a corpus. This process was conducted for the word “evaluation” in the CEJA2019, CAJA, and BNC to differentiate usage in academic evaluation, general academic, and general language.

Tables 10 through 12 portray the top ten thesaurus words from CEJA2019, CAJA, and BNC, respectively. “Evaluation” was entered as the search term. The “frequency” column displays the number of times a word appears in the corpus, and the “similarity” column displays “the percentage of collocates

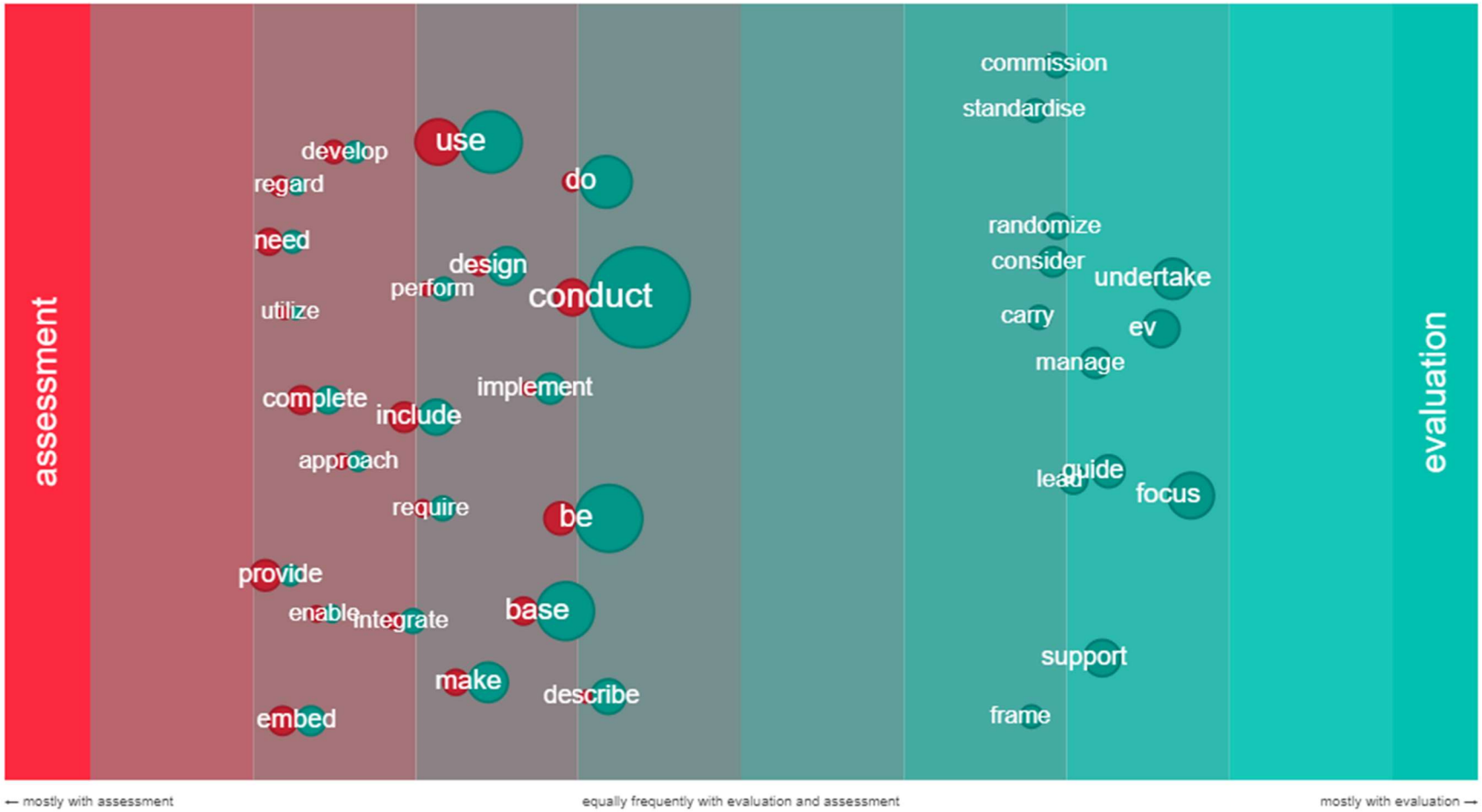

Figure 1. *Word Sketch Difference: Verbs with "Assessment" vs. "Evaluation" as object, CEJA2019 Corpus.*

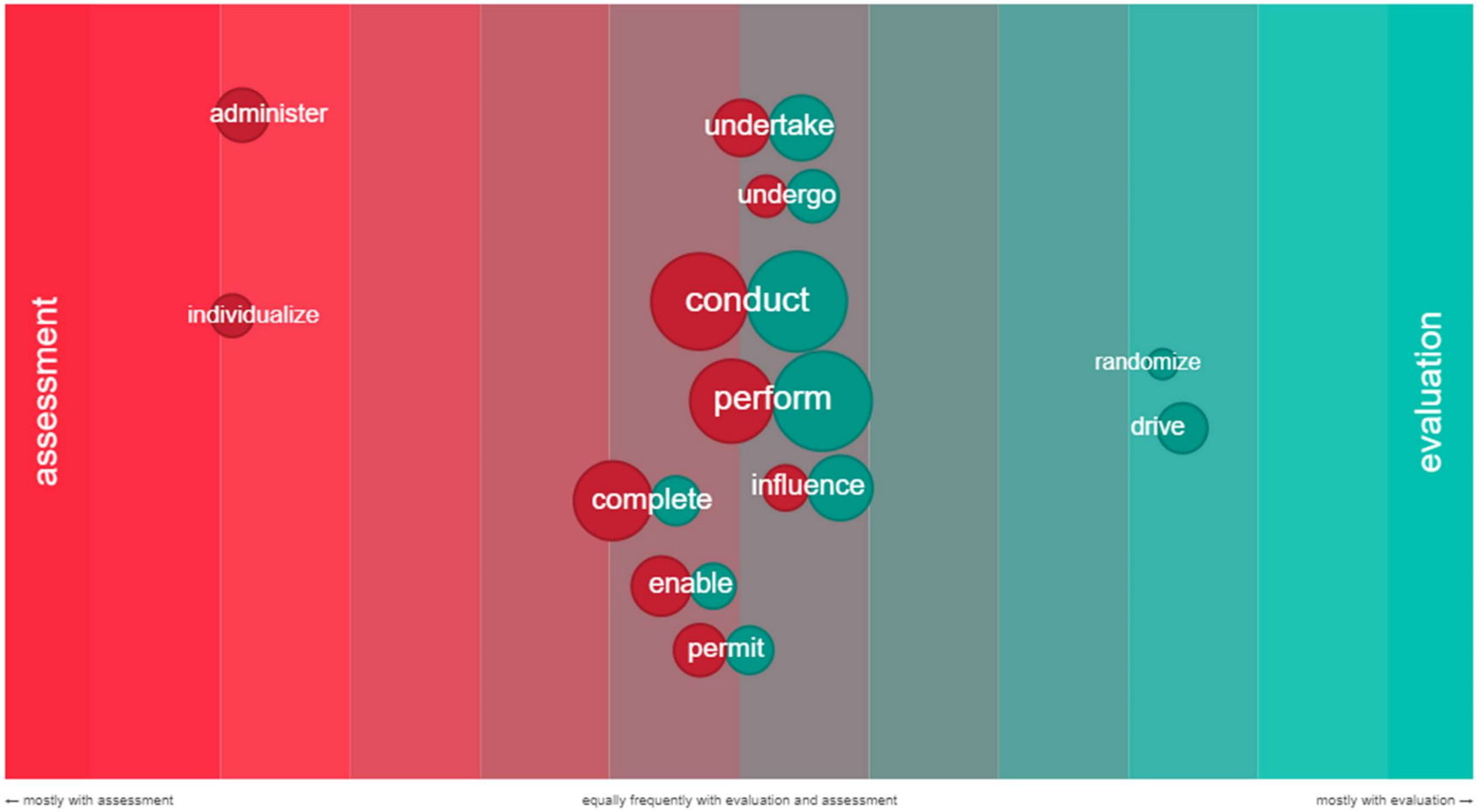

Figure 2. *Word Sketch Difference: Verbs with "Assessment" vs. "Evaluation" as object, CAJA Corpus.*

Table 10

Synonyms of “evaluation” from CEJA2019

Synonym	Similarity	Frequency
research	0.397	9,016
program	0.368	12,032
study	0.362	10,185
process	0.345	5,802
assessment	0.334	4,084
intervention	0.316	4,947
practice	0.313	4,199
development	0.311	3,954
project	0.302	4,998
approach	0.298	4,881

Table 11

Synonyms of “evaluation” from CAJA

Synonym	Similarity	Frequency
assessment	0.456	12,296
interpretation	0.34	16,149
selection	0.337	11,724
description	0.336	10,661
understanding	0.335	16,373
examination	0.334	7,076
perception	0.334	11,472
identification	0.329	8,776
investigation	0.325	8,798
design	0.322	19,939

Table 12

Synonyms of “evaluation” from BNC

Synonym	Similarity	Frequency
assessment	0.378	7,807
review	0.262	8,891
appraisal	0.26	1,180
planning	0.255	9,577
investigation	0.244	6,740
examination	0.242	6,186
testing	0.241	1,980
analysis	0.24	14,113
monitoring	0.232	1,667
interpretation	0.217	5,343

one uses the words “evaluation” and “assessment” interchangeably, they are likely to show up on the thesaurus list. The relative absence of words that seem like they could be confused for evaluation on the CEJA2019 thesaurus list suggests that evaluation scholars tend to use the word less liberally than the academy or general speakers.

Summary of Findings

This section described the results of the methods used to explore the language of evaluation scholarship. The results of the lexical diversity analysis, keyword procedures and coverage analyses all revealed a lexicon that was very large and diverse, but not especially technical. However, the thesaurus, word sketches, and word sketch differences revealed that while the words themselves may not be especially technical, the usage applied to certain keywords and phrases is unique. For example, differences in the thesaurus for CEJA2019, CAJA and BNC reveal that evaluation scholars may be more precise about their use of the

word “evaluation”. The word sketch differences reveal that evaluators are also less likely to use “evaluation” and “assessment” interchangeably than the academy in general.

CHAPTER V
DISCUSSION

This study was an exploratory inquiry employing the tools of corpus linguistics to illuminate the nature of scholarly evaluation language. This chapter focuses on the findings of this inquiry and the limitations and implications of the research. In addition, conjectures as to the meaning of what was learned and warranted conclusions for evaluation scholarship and practice are made, where applicable.

Summary of Results

RQ1: To what extent does the scholarly evaluation literature possess a unique lexicon? The investigations conducted in this study yield insights to this question in several different ways. First, the lemma list produced 70,755 unique lemmas in the corpus. Comparative coverage calculations reveal that a comparatively large number of unique lemmas needs to be added (4,000) to the New General Service List and New Academic Service List to achieve 95% coverage of the corpus, considered by some an essential benchmark for achieving readability (Coxhead, 2000). This is also suggestive of a lexicon that is very diverse and of low specification.

The keyword procedure produced a set of 412 key lemmas that do not appear on the New Academic Word List or the New General Service List. This demonstrates a distinct set of lexical items overrepresented in the scholarly evaluation literature compared to other general scholarship. This is an expected result. However, this list provides only limited coverage of the CEJA2019 Corpus compared to other specialized word lists. Moreover, this confirms the previous finding that the evaluation lexicon is diverse compared to other disciplines.

Examining the keywords from the Evaluation Specific Word List and the Evaluation Word List according to Jablonkai's (2010) levels of lexis revealed a

relatively low level of specialization, with only two technical terms ("evaluative" and "evaluand") among the top thirty words on the Evaluation Specific Word List.

Examining the top thirty *n*-grams also reveals few examples of technical lexis. While the *n*-grams reveal something as to the "aboutness" of evaluation literature, they do not suggest a highly specialized lexis consisting of many technical terms.

Relying simply on keywords and *n*-grams gives the sense that there is not much unique to evaluation language at this level. However, the word sketches, word sketch differences, and the thesaurus functions were used to examine *how*, rather than whether, specialized evaluation terminology is used. When comparing between various corpora using this tool, some specific patterns involving the use of the word "evaluation" emerge. Of particular interest are the modifiers of "evaluation" in CEJA2019 compared to CAJA; some evaluation-specific terminology emerged. These modifiers were recognizable as discipline-specific technical terms relating to the taxonomy of evaluation. This is where discipline-specific terms such as "empowerment evaluation" and "realist evaluation" were observed. This demonstrates the focus of evaluation literature on the theories, models and approaches of evaluation.

Thesaurus entries also revealed some differences in semantic equivalents between the CEJA2019, CAJA, and BNC corpora. The CEJA2019 thesaurus entries were far more likely to include words that are not conceptually adjacent or equivalent to "evaluation," for example, including words such as "project," "practice," and "intervention." General academic language and, to a greater extent, general English were more likely to include conceptually adjacent words such as "assessment" and "appraisal." This suggests that evaluation scholars use the word "evaluation" less interchangeably with other words. In other words, evaluation scholars do not use the word "evaluation" as haphazardly or imprecisely as others.

Upon examining the word sketch differences from the CEJA2019 corpus for the entries "evaluation" and "assessment," the results suggest that when

operating as the subject of a verb, where the word "assessment" appears in the evaluation literature, "evaluation" could cover almost any instance. However, the reverse is not true, meaning that "assessment" cannot cover almost any instance of the word "evaluation". When examining the results of the same operation conducted in the CAJA corpus, "evaluation" and "assessment" appear to be almost universally interchangeable. This suggests again that the word "evaluation" is used in some very specialized and precise ways in the evaluation literature compared to other terms and their use in other corpora.

All of these things considered together suggest that the evaluation lexis is relatively diverse and unspecialized. However, findings also suggest that although the specialized lexis of the evaluation literature is limited, the ways in which evaluation scholars use evaluation terms is more specialized than general academic or general language. The results also reveal the potential "aboutness" of the literature: the discussion of evaluation theories, models, and approaches.

RQ2: What are the most significant keywords in the evaluation lexicon? This research question is relatively simple to answer. In essence, the answer is the product of the keyword procedures described in Chapter III. These results, the Evaluation Word LIst and Evaluation Specific Word List, which are displayed in their entirety in Appendices B and C, help identify the most common terms in the evaluation literature and those considered evaluation-specific due to their absence from the New General Service List and the New Academic Word List. The resulting lists may have applications for those seeking to develop pedagogical resources in evaluation practice and scholarship.

Aside from these initial research questions, several other secondary questions (SQ) that could be answered arose during the examination of the results. These items and the results suggesting their answers are briefly described below.

Interpretations

These results are aligned with some fundamental issues surrounding the nature of evaluation and the state of the field. The first of these issues is the extent to which evaluation has achieved a state of unification similar to other fields. Schwandt (2015) points out that despite some level of variability in fields such as medicine, law, and accounting, practitioners in these fields "share something of a common conception of the social purpose of their respective undertakings as well as a professional identity" (p. 27). Some assert that a common identity or purpose continues to evade the evaluation community. Related to this is the question of whether a distinct discourse community of evaluation scholars exists. Swales (1991, 2016) pointed to the need for a commonly-defined purpose—or one that at least could conceivably be discovered.

There is a small set of evaluation-specific lexis that is germane to evaluation. This is mainly related to evaluation taxonomy, specifically names for theories, models, and approaches of evaluation. Some may see these pieces of evidence as evidence against the status of evaluation scholars as a discourse community, given the low level of lexical complexity demonstrated. However, this fixation on the unspecialized nature of the individual lexical items that appear on the keyword lists would miss the larger point. In these findings we see that a central concern of evaluation scholarly literature does indeed emerge: The focus on evaluation theory, specifically. We also observe that evaluation scholars are more precise with their definitions of evaluation and are less apt to use the term interchangeably than other scholars or general English speakers.

Implications

As William Shadish stated in his (1998) American Evaluation Association address, "if you do not know much about evaluation theory, you are not an evaluator" (p. 6). Indeed, as the title of Shadish's speech asserts, these results seem to confirm that "evaluation theory is who we are." These findings demonstrate the centrality of evaluation theory in the evaluation literature, bearing evidence to both add considerable weight to this assertion, and to calm the fears of some who worry that this core subject area (Scriven 1993) may be diluted in evaluation scholarship.

This study may also present some evidence to allay the concerns of those who fear that "fuzzy" definitions of evaluation are rampant in the field. This evidence suggests that evaluation scholars at least have a more precise definition of evaluation than other scholars or the general public. Whether the integrity of these definitions are up to the standards of those who voice this concern is another issue and warrants further investigation. Additionally, Christie and Rose's (2003) fear that these unclear definitions could translate to poor practice is a conjecture that would need to be tested.

The findings related to a lack of highly specialized vocabulary could signal the need for continued professionalization of evaluation scholars and practitioners. One might expect to have a more cohesive core terminology developed at this point. That this was not detected could be due to such terminology not existing or simply that it is there but has not been adopted. Given resources in the field such as the *Evaluation Thesaurus* (Scriven, 1991) and *Encyclopedia of Evaluation* (Mathison, 2005), it seems likely that it is there but has not been adopted widely by scholars. In either case—adoption or development—it would seem that more engagement in evaluation-specific pedagogy and more scholarship in research on evaluation, both through a credentialing requirement and the development of more evaluation academic programs, could serve both purposes.

In addition to increasing evaluation pedagogy and scholarly involvement in evaluation research and theory development, increased engagement with the evaluation literature in fields that use evaluation and its methods could benefit the development of the evaluation lexicon. Disciplines such as international development, social work, and economics, for example, each have their own evaluation literature. Those who are creating this material are also advised to reference and direct students towards foundational scholarly sources in the evaluation literature. This could aid in bolstering evaluation as a united transdiscipline.

Limitations

This is an exploratory and novel study in the field of evaluation. As such, it makes this work both exciting and challenging, as there is little to compare it to that could lend credibility to these results, other than the integrity of the methods themselves. Described below are the various challenges associated with the investigation.

Since there is a heavy emphasis on comparison with other corpora, it is impossible to know the extent to which corpora share certain similarities. For instance, some of the reference corpora used in this study were produced with a significant amount of computing power and programming expertise. This could have resulted in corpora that are much cleaner than the current corpus. Or, less clean. It is unknown, and this limits any conclusions as to the difference between the study corpus and reference corpora.

Second, because of the highly specific nature of corpora, the use of a range of different programs and operations, and different conditions surrounding this and other corpus linguistic studies cited in this work, it can be difficult to make precise comparisons. Any conjectures about the language of evaluation in reference to other professions are tenuous and warrant further investigation.

Third, because of limited data availability in terms of the reference corpora, this limited the available tools for this study and, ultimately, the researcher's choices.

Fourth are limitations related to the design of the focus corpus. The decision to bound the timeframe to the year 2019 restricts the extent to which one might be able to make inferences to past or future scholarship. The focus corpus also includes only one genre, making it difficult to speak to similar issues in the broader evaluation community.

Lastly, this study, being exploratory in nature, begins to offer just a glimpse of what diving into linguistic inquiry in evaluation literature may offer. The thesaurus, word sketch, and word sketch differences tools could be used to explore many more words and many different grammatical associations that were outside of the scope of this study. Therefore, these findings are extremely tentative and there is ample room for refutation or contradiction. For example, use of the word sketch tool to investigate the usage of "evaluation" in the literature explored one small aspect of the issue. Further investigation may support or refute the findings derived in this study.

Opportunities for Future Research

This work follows a line of research on evaluation scholarship and serves to further this cause. Additionally, it opens up an entirely new line of inquiry. This is the first empirical study investigating the language of evaluation scholarship using corpus linguistics or any other quantitative method of inquiry.

The development of CEJA2019, Evaluation Word LIst, and Evaluation Specific Word List represents a cornerstone for linguistic inquiry in evaluation scholarship. Now that this groundwork for beginning to understand the size, diversity, and discipline-specificity of evaluation scholarship language has been laid, several other questions are raised. For example, how exactly does the language of evaluation compare to other related fields such as sociology,

organizational psychology, nonprofit management, and public health? Answering this question using corpus-based methods, and CEJA2019 specifically, could yield insights regarding the relatedness of evaluation to other fields, such as the relatedness of the subfields as identified in Gilmore and Millar (2018). Additionally, more inquiries as to the usage of specific lexical items (e.g., what differentiates the use of “evaluation” from “research”, or, from “science”?) are now possible. There could also be opportunities to evaluate the utility of the Evaluation Specific Word List in terms of evaluation pedagogy and in terms of its coverage of other types of evaluation literature. More work could also be done to expand the CEJA series beyond CEJA2019 to monitor any changes in language use over time or build a broader, more generalizable corpus.

Within CEJA there are also opportunities for comparing subcorpora to explore various issues of linguistic diversity among evaluation scholars. For instance, the geographical classifications in the results section could facilitate the comparisons of evaluation language between North America and Europe, for instance. Also possible would be the comparison of language between the interdisciplinary journals such as the American Journal of Evaluation and more discipline-specific journals such as Evaluation in the Health Professions or Studies in Educational Evaluation.

Aside from research using CEJA or the Evaluation Specific Word List, a parallel inquiry is also possible using corpus linguistics to examine non-scholarly evaluation writing or speech. Analysis of evaluation reports, conference proceedings, correspondence, or blogs could reveal findings that are relevant to professional evaluation practice, pedagogy, community, and so forth. This work could supplement the CEJA series, working toward constructing a general evaluation corpus that includes more than just academic publications. Such corpora could support studies of how language varies across evaluation approaches, or investigage issues of how language influences practice as suggested by Christie and Rose (2003).

REFERENCES

Baker, P., Hardie, A., & McEnery, T. (2006). *A glossary of corpus linguistics*. Edinburgh: Edinburgh University Press.

Bancroft-Billings, S. (2020). Identifying spoken technical legal vocabulary in a law school classroom. *English for Specific Purposes*, *60*, 9–25. https://doi.org/10.1016/j.esp.2020.04.003

Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure*. Amsterdam: John Benjamins Publishing Co.

Borg, E. (2003). Discourse community. *English Language Teaching*, *57*(4), 398–400.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, *36*(1), 1–22. https://doi.org/10.1093/applin/amt018

Brezina, V., & Gablasova, D. (2017). How to Produce Vocabulary Lists? Issues of Definition, Selection and Pedagogical Aims. A Response to Gabriele Stein. *Applied Linguistics*, *38*(5), 764–767. https://doi.org/10.1093/applin/amx022

Brown, C. L. (2000). Sociolinguistic dynamics of gender in focus groups. *New Directions for Evaluation*, *2000*(86), 55–67. https://doi.org/10.1002/ev.1172

Browne, C., Culligan, B., & Phillips, J. (2013). The new general service list. Retrieved September 30, 2020, from http://newgeneralservicelist.org

Cabatoff, K. (2000). Translating evaluation findings into "policy language." *New Directions for Evaluation*, *2000*(86), 43–54. https://doi.org/10.1002/ev.1171

CDC. (2011). Self-Study Guide - Program Evaluation. *Program Performance and Evaluation Office (PPEO)*, (October), 103. Retrieved from https://www.cdc.gov/eval/guide/index.htm

Cheng, W. (2011). *Exploring corpus linguistics: Language in action*. London: Routledge.

Chiba, Y., Millar, N., & Budgell, B. (2010). The language of midwifery and perinatal care: a quantitative analysis. *Journal of Japan Academy of Midwifery*, *24*(1), 74–83. https://doi.org/10.3418/jjam.24.74

Christie, C. A. (2013). Michael Scriven and the Development of the Evaluation Lexicon. In S. I. Donaldson (Ed.), *The Future of Evaluation in Society: A Tribute to Michael Scriven* (pp. 93–105). Charlotte, NC: Information Age Publishing.

Christie, C. A., Lemire, S., & Inkelas, M. (2017). Understanding the Similarities and Distinctions Between Improvement Science and Evaluation. *New Directions for Evaluation*, *2017*(153), 11–21. https://doi.org/10.1002/ev.20237

Christie, C., & Rose, M. (2003). The language of evaluation theory: Insights gained from an empirical study of evaluation theory and practice. *The Canadian Journal of Program Evaluation*, *18*(2), 33.

Coryn, C.L.S., & Hattie, J. (2006). The Transdisciplinary Model of Evaluation. *Journal of Multidisciplinary Evaluation*, *3*(4), 107–114.

Coryn, Chris L S, Noakes, L. A., Westine, C. D., & Schroter, D. C. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. https://doi.org/10.1177/1098214010389321

Coryn, Chris L S, Wilson, L. N., Westine, C. D., Hobson, K. A., Ozeki, S., Fiekowsky, E. L., … Schröter, D. C. (2017). A Decade of Research on Evaluation. *American Journal of Evaluation*, *38*(3), 329–347. https://doi.org/10.1177/1098214016688556

Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. *English for Specific Purposes*, *51*, 84–97. https://doi.org/10.1016/j.esp.2018.03.006

Dahler-Larsen, P., Abma, T., Bustelo, M., Irimia, R., Kosunen, S., Kravchuk, I., … Tshali, C. K. (2017). Evaluation, Language, and Untranslatables.

American Journal of Evaluation, *38*(1), 114–125. https://doi.org/10.1177/1098214016678682

Education Endowment Foundation. (n.d.). Evaluation Glossary. Retrieved April 2, 2021, from https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/evaluation-information-for-grantees/glossary/

Gilmore, A., & Millar, N. (2018). The language of civil engineering research articles: A corpus-based approach. *English for Specific Purposes*, *51*, 1–17. https://doi.org/10.1016/j.esp.2018.02.002

Grabowski, Ł. (2015). Keywords and lexical bundles within english pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, *38*, 23–33. https://doi.org/10.1016/j.esp.2014.10.004

Jablonkai, R. R. (2010). *A corpus-linguistic investigation into the lexis of written English EU discourse: An ESP pedagogic perspective*. Eotvos Lorand University.

Jackson, H. (2002). The beginnings. In *Lexicography: An introduction* (pp. 31–46). London: Routledge.

Kaminsky, A. (2000). Beyond the literal: Metaphors and why they matter. *New Directions for Evaluation*, *2000*(86), 69–80. https://doi.org/10.1002/ev.1173

Kilgarriff, A. (2009). Simple maths for keywords. In *Proceedings of the Corpus Linguistics Conference. Liverpool, UK. 2009.* (p. 6).

Kosem, I. (2010). *Designing a model for a corpus-driven dictionary of academic English*. Retrieved from https://publications.aston.ac.uk/id/eprint/14664/

Madison, A. M. (2000). Language in defining social problems and in evaluating social programs. *New Directions for Evaluation*, *2000*(86), 17–28. https://doi.org/10.1002/ev.1169

Mathison, S. (2005). *Encyclopedia of evaluation*. Thousand Oak, CA: Sage publications.

National Endowment for Financial Education. (n.d.). Appendix A: Glossary of evaluation terms. Retrieved April 2, 2021, from https://toolkit.nefe.org/evaluation-resources/evaluation-manual/section-5-appendices/appendix-a-glossary-of-evaluation-terms

Nguyen Le, C. N., & Miller, J. (2020). A corpus-based list of commonly used English medical morphemes for students learning English for specific purposes. *English for Specific Purposes*, *58*, 102–121. https://doi.org/10.1016/j.esp.2020.01.004

Organisation for Economic Co-operation and Development. (2002). Glossary of key terms in evaluation and results based management. https://doi.org/10.1787/9789264034921-en-fr

Patton, M. (2000). Overview : Language Matters. *New Directions for Evaluation*, *2000*(86), 5–16.

Patton, M. Q. (2018). Evaluation Science. *American Journal of Evaluation*, *39*(2), 183–200. https://doi.org/10.1177/1098214018763121

Pearson, J. (1998a). 0 Introduction. In J. Pearson (Ed.), *Terms in context* (pp. 1–6). Amsterdam: John Benjamins Publishing Co. https://doi.org/10.30687/978-88-6969-232-1/001

Pearson, J. (1998b). Corpora, corpus design and corpus selection. In *Terms in context* (pp. 41–66).

Pearson, J. (1998c). Identifying differences between words and terms. In *Terms in context*.

Rayson, P., Berridge, D., & Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. *JADT 2004 : 7es Journées Internationales d'Analyse Statistique Des Données Textuelles*, 1--12. Retrieved from http://eprints.lancs.ac.uk/12424/

Rea-Rizzo, C. (2010). Getting on with corpus compilation: From theory to practice. *ESP World*, *9*(1(27)), 1–23. Retrieved from http://www.esp-world.com/Articles_27/Camino Rea.pdf

Sager, J. C. (1990). *A practical course in terminology processing*. Amsterdam:

John Benjamins Publishing Co.

Schiffrin, D. (2011). Discourse analysis (linguistic). In P. C. Hogan (Ed.), *The Cambridge encyclopedia of the language sciences*. Cambridge, UK: Cambridge University Press. Retrieved from http://libproxy.library.wmich.edu/login?url=https://search.credoreference.com/content/entry/cupelanscis/discourse_analysis_linguistic/0?institutionId=1566

Schwandt, T. A. (2015). *Evaluation foundations revisited: Cultivating a life of the mind for practice*. Stanford: Stanford University Press.

Scott, M. (2021). Personal Correspondence, February 9.

Scott, M., & Tribble, C. (2006). *Textual patterns : Key words and corpus analysis in language education*. Amsterdam: John Benjamins Publishing Co.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: SAGE.

Scriven, M. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*, *20*(1), 147–166. https://doi.org/10.1080/17474124.2017.1343144

Scriven, M. (2013). The Foundation and Future of Evaluation. In S. I. Donaldson (Ed.), *The Future of Evaluation in Society: A Tribute to Michael Scriven* (pp. 11–44). Charlotte, NC: Information Age Publishing.

Shadish, W. R. (1998). Evaluation theory is who we are. *American Journal of Evaluation*, *19*(1), 1–19. https://doi.org/10.1177/109821409801900102

Shanker, S. (2019). *Definitional tension: The cunstruction of race in and through evaluation*. University of Minnesota.

Sinclair, J. (2005). Corpus and text: Basic principles. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice* (pp. 5–24). Oxford: AHDS Literature, Languages and Linguistics. Retrieved from http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm

Sketch Engine: Thesaurus. (n.d.). Retrieved June 14, 2021, from https://app.sketchengine.eu/#thesaurus?corpname=user%2Fawkates%2F

eval_rsx2019&tab=about&lemma=evaluation&lpos=-n&showresults=1

Smith, N. L. (1979). Evaluation reflections: Requirements for a discipline of evaluation. *Studies In Educational Evaluation*, *5*, 5–12. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/0191491X81900018

Sportanddev.org. (n.d.). Glossary of key terms in monitoring and evaluation. Retrieved April 2, 2021, from https://www.sportanddev.org/en/toolkit/monitoring-and-evaluation/glossary-key-terms-monitoring-and-evaluation

Staples, S. (2019). Using corpus-based discourse analysis for curriculum development: Creating and evaluating a pronunciation course for internationally educated nurses. *English for Specific Purposes*, *53*, 13–29. https://doi.org/10.1016/j.esp.2018.08.005

Stubbs, M. (2010). Three concepts of keywords. In M. Scott & M. Bondi (Eds.), *Keyness in texts* (pp. 21–42). Amsterdam: John Benjamins Publishing Co. https://doi.org/10.1075/scl.41.03stu

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, & applications* (2nd ed.). San Francisco: Jossey-Bass.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, J. M. (2016). Reflections on the concept of discourse community. *La Revue de GERAS*, *69*, 7–19. https://doi.org/10.4000/asp.4774

Switchboard. (n.d.). Monitoring and Evaluation Glossary. Retrieved February 2, 2021, from https://switchboardta.org/resource/monitoring-and-evaluation-glossary/

US Department of Energy. (n.d.). Program evaluation: Glossary. Retrieved April 2, 2021, from https://www.energy.gov/eere/analysis/program-evaluation-glossary

US Fish and Wildlife Service. (n.d.). Glossary. Retrieved April 2, 2021, from https://www.fws.gov/urban/evaluation/glossary

Wang, S., & Huang, C. (2017). Word sketch lexicography : new perspectives on

lexicographic studies of Chinese near synonyms. *Lingua Sinica*, *3*(11), 1–22. https://doi.org/10.1186/s40655-017-0025-4

Appendix A

Articles Included in CEJA2019

Abd El Razik, M. S., & Rashad Salem, M. (2019). From public health and demographic research to decision making: An intervention study in Giza Governorate-Egypt. *Evaluation and Program Planning*, *77*(August), 101704. https://doi.org/10.1016/j.evalprogplan.2019.101704

Abma, T. A. (2019). Book review. *Evaluation and Program Planning*, *74*, 18–19. https://doi.org/10.1016/j.evalprogplan.2019.02.002

Abrahams, M. A. (2019). Editorial-2019: Omniscience of monitoring and evaluation. *African Evaluation Journal*, *7*(1), 1–2. https://doi.org/10.4102/AEJ.V7I1.433

Acree, J. (2019). Applying the collaborative approaches to Evaluation (CAE) principles in an educational evaluation: Reflections from the field. *Canadian Journal of Program Evaluation*, *34*(2), 272–281. https://doi.org/10.3138/cjpe.53185

Ahmad, M., & Jan, M. A. (2019). Diversity of information sources: An evaluation of global think tanks knowledge construct. *Research Evaluation*, *28*(3), 273–278. https://doi.org/10.1093/reseval/rvz011

Alawode, O. O., & Oluwatayo, I. B. (2019). Development Outcomes of Fadama III among Fish Farmers in Nigeria: Evidence from Lagos State. *Evaluation and Program Planning*, *75*(February), 10–19. https://doi.org/10.1016/j.evalprogplan.2019.02.004

Alderman, L. (2019). Evaluation Perspectives. *Evaluation Journal of Australasia*, *19*(1), 49–53. https://doi.org/10.1177/1035719x19837025

Alindogan, M. A. (2019). Evaluation competencies and functions in advertised evaluation roles in Australia. *Evaluation Journal of Australasia*, *19*(2), 88–100. https://doi.org/10.1177/1035719X19857197

Alkin, M. C., & Christie, C. A. (2019). Theorists' Models in Action: A Second Look. *New Directions for Evaluation*, *2019*(163), 11–18. https://doi.org/10.1002/ev.20377

Allen, S., & Peterman, K. (2019). Evaluating Informal STEM Education: Issues and Challenges in Context. *New Directions for Evaluation*, *2019*(161), 17–33. https://doi.org/10.1002/ev.20354

Alnahdi, G. H. (2019). Rasch validation of the Arabic version of the Teacher Efficacy for Inclusive Practices (TEIP)scale. *Studies in Educational Evaluation*, *62*(June), 104–110. https://doi.org/10.1016/j.stueduc.2019.05.004

Alotaibi, B. A., Yoder, E., Brennan, M. A., & Kassem, H. S. (2019). Training needs of extension agents' regarding organic agriculture in Saudi Arabia. *Evaluation and Program Planning*, *77*(March). https://doi.org/10.1016/j.evalprogplan.2019.101711

Al-Qatawneh, S., & Al Rawashdeh, A. (2019). Gender representation in the Arabic language textbook for the ninth grade approved by the Ministry of Education for use in schools in the United Arab Emirates (UAE). *Studies in Educational Evaluation*, *60*(June 2018), 90–98. https://doi.org/10.1016/j.stueduc.2018.12.001

Altindag, O., Joyce, T. J., & Reeder, J. A. (2019). Can Nonexperimental Methods Provide Unbiased Estimates of a Breastfeeding Intervention? A Within-Study Comparison of Peer Counseling in Oregon. *Evaluation Review*, *43*(3–4), 152–188. https://doi.org/10.1177/0193841X19865963

Alvarez, S., Bampasidou, M., & Solís, D. (2019). Evaluating the Impact of Employing Local Tax Collectors to Improve State-Level Licensing Program Outcomes in Florida. *Evaluation Review*, *43*(1–2), 77–107. https://doi.org/10.1177/0193841X19865353

Amundsen, C., & D'Amico, L. (2019). Using Theory of Change to evaluate socially-situated, inquiry-based academic professional development. *Studies in Educational Evaluation*, *61*(September 2018), 196–208. https://doi.org/10.1016/j.stueduc.2019.04.002

Andersson, U. B., Löfgren, H., & Gustafson, S. (2019). Forward-looking assessments that support students' learning: A comparative analysis of two approaches. *Studies in Educational Evaluation*, *60*(August 2018), 109–116. https://doi.org/10.1016/j.stueduc.2018.12.003

Andrews, N. C. Z., Pepler, D. J., & Motz, M. (2019). Research and Evaluation With Community-Based Projects: Approaches, Considerations, and Strategies. *American Journal of Evaluation*, *40*(4), 548–561. https://doi.org/10.1177/1098214019835821

Annesi, J. J. (2019). Generalizability of an adult-validated model for predicting increased physical activity in after-school program children: Effects of treatment formats on fitness indicators. *Evaluation and Program Planning*, *73*(August 2018), 33–43. https://doi.org/10.1016/j.evalprogplan.2018.11.008

Annesi, J. J. (2019). Change in behavioral exercise program-associated self-regulation enhances self-regulation-induced eating improvements across levels of obesity severity. *Evaluation and Program Planning*, *75*(March), 31–37. https://doi.org/10.1016/j.evalprogplan.2019.04.002

Arteaga, I., Thornburg, K., Darolia, R., & Hawks, J. (2019). Improving Teacher Practices With Children Under Five: Experimental Evidence From the Mississippi Buildings Blocks. *Evaluation Review*, *43*(1–2), 41–76. https://doi.org/10.1177/0193841X19865070

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in multiple regression analysis: What it is, how to detect it and how to solve it with applications in R and SPSS. *Practical Assessment, Research and Evaluation*, *24*(1), 1–16.

Atinga, R. A., Agyepong, I. A., & Esena, R. K. (2019). Willing but unable? Extending theory to investigate community capacity to participate in Ghana's community-based health planning and service implementation. *Evaluation and Program Planning*, *72*(July 2018), 170–178. https://doi.org/10.1016/j.evalprogplan.2018.10.001

Atteberry, A., Bassok, D., & Wong, V. C. (2019). The Effects of Full-Day Prekindergarten: Experimental Evidence of Impacts on Children's School Readiness. *Educational Evaluation and Policy Analysis*, *41*(4), 537–562. https://doi.org/10.3102/0162373719872197

Auger, M., Crooks, C. V., Lapp, A., Tsuruda, S., Caron, C., Rogers, B. J., & van der Woerd, K. (2019). The essential role of cultural safety in developing culturally-relevant prevention programming in First Nations communities: Lessons learned from a national evaluation of Mental Health First Aid First Nations. *Evaluation and Program Planning*, *72*(December 2017), 188–196. https://doi.org/10.1016/j.evalprogplan.2018.10.016

Baffoe, G. (2019). Exploring the utility of Analytic Hierarchy Process (AHP) in ranking livelihood activities for effective and sustainable rural development interventions in developing countries. *Evaluation and Program Planning*, *72*(March 2018), 197–204. https://doi.org/10.1016/j.evalprogplan.2018.10.017

Baldwin, C. K., & Lander, R. (2019). Developmental Evaluator Functional Role Activities and Programmatic Developments: A Case Study Analysis. *American Journal of Evaluation*, *40*(1), 35–54. https://doi.org/10.1177/1098214017743586

Banke-Thomas, A., Nieuwenhuis, S., Ologun, A., Mortimore, G., & Mpakateni, M. (2019). Embedding value-for-money in practice: A case study of a health pooled fund programme

implemented in conflict-affected South Sudan. *Evaluation and Program Planning*, *77*(October). https://doi.org/10.1016/j.evalprogplan.2019.101725

Barnow, B. S., & Greenberg, D. H. (2019). Special Issue Editors' Essay. *Evaluation Review*, *43*(5), 231–265. https://doi.org/10.1177/0193841X19865076

Barri, R. (2019). Les indicateurs sont morts, vive les indicateurs! Towards a political economy of S&T indicators: A critical overview of the past 35 years. *Research Evaluation*, *28*(1), 2–6. https://doi.org/10.1093/reseval/rvy029

Barrington, G. V. (2019). Consulting After 50: Redirection and Reinvention for Career Evaluators. *New Directions for Evaluation*, *2019*(164), 155–167. https://doi.org/10.1002/ev.20391

Bartanen, B., Grissom, J. A., & Rogers, L. K. (2019). The Impacts of Principal Turnover. *Educational Evaluation and Policy Analysis*, *41*(3), 350–374. https://doi.org/10.3102/0162373719855044

Bartels, S. A., Michael, S., Vahedi, L., Collier, A., Kelly, J., Davison, C., … Geara, P. (2019). SenseMaker® as a monitoring and evaluation tool to provide new insights on gender-based violence programs and services in Lebanon. *Evaluation and Program Planning*, *77*(June), 101715. https://doi.org/10.1016/j.evalprogplan.2019.101715

Bashkov, B. M., & Clauser, J. C. (2019). Determining item screening criteria using cost-benefit analysis. *Practical Assessment, Research and Evaluation*, *24*(2), 1–9.

Basileo, L. D., & Toth, M. (2019). A state level analysis of the Marzano Teacher Evaluation Model: Predicting teacher value-added measures with observation scores. *Practical Assessment, Research and Evaluation*, *24*(6). https://doi.org/10.7275/cc5b-6j43

Bast, L. S., Andersen, A., Ersbøll, A. K., & Due, P. (2019). Implementation fidelity and adolescent smoking: The X:IT study—A school randomized smoking prevention trial. *Evaluation and Program Planning*, *72*(January 2018), 24–32. https://doi.org/10.1016/j.evalprogplan.2018.09.004

Battistone, W., Buckmiller, T., & Peters, R. (2019). Assessing Assessment Literacy: Are new teachers prepared to assume jobs in school districts engaging in grading and assessment reform efforts? *Studies in Educational Evaluation*, *62*(April), 10–17. https://doi.org/10.1016/j.stueduc.2019.04.009

Bautista-Puig, N., García-Zorita, C., & Mauleón, E. (2019). European Research Council: Excellence and leadership over time from a gender perspective. *Research Evaluation*, *28*(4), 370–382. https://doi.org/10.1093/reseval/rvz023

Belcher, B. M., Ramirez, L. F., Davel, R., & Claus, R. (2019). A response to Hansson and Polk (2018) "assessing the impact of transdisciplinary research: The usefulness of relevance, credibility, and legitimacy for understanding the link between process and impact. *Research Evaluation*, *28*(2), 196–201. https://doi.org/10.1093/reseval/rvy037

Belfield, C. R., Bowden, A. B., & Rodriguez, V. (2019). Evaluating Regulatory Impact Assessments in Education Policy. *American Journal of Evaluation*, *40*(3), 335–353. https://doi.org/10.1177/1098214018785463

Bequette, M., Cardiel, C. L. B., Cohn, S., Kollmann, E. K., & Lawrenz, F. (2019). Evaluation Capacity Building for Informal STEM Education: Working for Success Across the Field. *New Directions for Evaluation*, *2019*(161), 107–123. https://doi.org/10.1002/ev.20351

Bhawra, J. (2019). Developing Monitoring and Evaluation Frameworks, by Anne Markiewicz and Ian Patrick. *Canadian Journal of Program Evaluation*, *34*(1), 165–167. https://doi.org/10.3138/cjpe.53169

Bilodeau, A., & Kranias, G. (2019). Self-evaluation tool for action in partnership: Translation and cultural adaptation of the original Quebec French tool to Canadian English. *Canadian Journal of Program Evaluation*, *34*(2), 192–206. https://doi.org/10.3138/cjpe.43685

Birch, L., & Jacob, S. (2019). “Deliverology” and evaluation: A tale of two worlds. *Canadian Journal of Program Evaluation*, *34*(2), 303–328. https://doi.org/10.3138/cjpe.53365

Blaser Mapitsa, C., & Chirau, T. J. (2019). Institutionalising the evaluation function: A South African study of impartiality, use and cost. *Evaluation and Program Planning*, *75*(February), 38–42. https://doi.org/10.1016/j.evalprogplan.2019.04.005

Blundo-Canto, G., Triomphe, B., Faure, G., Barret, D., De Romemont, A., & Hainzelin, E. (2019). Building a culture of impact in an international agricultural research organization: Process and reflective learning. *Research Evaluation*, *28*(2), 136–144. https://doi.org/10.1093/reseval/rvy033

Boadu, E. S., & Ile, I. (2019). Between power and perception: Understanding youth perspectives in participatory monitoring and evaluation (PM&E) in Ghana. *Evaluation and Program Planning*, *77*(October 2018), 101683. https://doi.org/10.1016/j.evalprogplan.2019.101683

Boehmer, H.-M., & Zaytsev, Y. K. (2019). Raising Aid Efficiency with International Development Aid Monitoring and Evaluation Systems. *Journal of MultiDisciplinary Evaluation*, *15*(32), 28–40. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=136082654&site=ehost-live

Bolhuis, E. (2019). The development of data use, data skills, and positive attitude towards data use in a data team intervention for teacher educators. *Studies in Educational Evaluation*, *60*(December 2018), 99–108. https://doi.org/10.1016/j.stueduc.2018.12.002

Borowiec, K., & College, B. (2019). Using rater cognition to improve generalizability of an assessment of scientific argumentation. *Practical Assessment, Research and Evaluation*, *24*(6). https://doi.org/10.7275/ey9d-p954

Boruch, R., Allen-Platt, C., & Gerstner, C. C. (2019). To Randomize or Not to Randomize? That is the Question. *New Directions for Evaluation*, *2019*(163), 73–82. https://doi.org/10.1002/ev.20373

Boshoff, N., & Sefatsa, M. (2019). Creating research impact through the productive interactions of an individual: An example from South African research on maritime piracy. *Research Evaluation*, *28*(2), 145–157. https://doi.org/10.1093/reseval/rvz001

Botha, L., Grobbelaar, S. S., & Bam, W. G. (2019). Developing an evaluation framework for university-driven technology-based, innovation for inclusive development (UTI4ID) projects. *Research Evaluation*, *28*(4), 326–343. https://doi.org/10.1093/reseval/rvz021

Boumans, D., & Ferry, M. (2019). Networks and efficient policy implementation: Insights from cohesion policy. *Evaluation*, *25*(4), 411–429. https://doi.org/10.1177/1356389019855096

Bourgeois, I. (2019). Editor's remarks. *Canadian Journal of Program Evaluation*, *34*(1), V–VI. https://doi.org/10.3138/cjpe.34.1.v

Bourgeois, I., & Buetti, D. (2019). Using Action Research to Build Evaluation Capacity in Public Health Organizations. *Journal of MultiDisciplinary Evaluation*, *15*(33), 81–90.

Bowman, N. (2019). Nation-to-nation evaluation: Governance, tribal sovereignty, and systems thinking through culturally responsive indigenous evaluations. *Canadian Journal of Program Evaluation*, *34*(2), 343–356. https://doi.org/10.3138/cjpe.67977

Boyce, A. S. (2019). Introduction to section on values in evaluation praxis. *Evaluation Journal of Australasia*, *19*(4), 195–198. https://doi.org/10.1177/1035719X19892526

Boyce, A. S., Avent, C., Adetogun, A., Servance, L., DeStefano, L., Nerem, R., & Platt, M. O. (2019). Implementation and evaluation of a biotechnology research experience for African-American high school students. *Evaluation and Program Planning*, *72*(September 2018), 162–169. https://doi.org/10.1016/j.evalprogplan.2018.10.004

Boyce, A. S., & McGowan, B. L. (2019). An Exploration of Two Novice Evaluation Educators' Experiences Developing and Implementing Introduction to Evaluation Courses. *American Journal of Evaluation*, *40*(1), 119–136. https://doi.org/10.1177/1098214018778812

Brand, S. L., Quinn, C., Pearson, M., Lennox, C., Owens, C., Kirkpatrick, T., … Byng, R. (2019). Building programme theory to develop more adaptable and scalable complex interventions: Realist formative process evaluation prior to full trial. *Evaluation*, *25*(2), 149–170. https://doi.org/10.1177/1356389018802134

Bremner, L. (2019). Creating new stories: The role of evaluation in truth and reconciliation. *Canadian Journal of Program Evaluation*, *34*(2), 331–342. https://doi.org/10.3138/cjpe.67976

Briggs, L., Trautmann, N., & Phillips, T. (2019). Exploring challenges and lessons learned in cross-cultural environmental education research. *Evaluation and Program Planning*, *73*(September 2018), 156–162. https://doi.org/10.1016/j.evalprogplan.2019.01.001

Brook, J., & Akin, B. (2019). Using theory of change as a framework for examining community context and philanthropic impact. *Evaluation and Program Planning*, *77*(August), 101708. https://doi.org/10.1016/j.evalprogplan.2019.101708

Brunner, E., Cowen, J. M., Strunk, K. O., & Drake, S. (2019). Teacher Labor Market Responses to Statewide Reform: Evidence From Michigan. *Educational Evaluation and Policy Analysis*, *41*(4), 403–425. https://doi.org/10.3102/0162373719858997

Brunner, R., Craig, P., & Watson, N. (2019). Evaluability assessment: An application in a complex community improvement setting. *Evaluation*, *25*(3), 349–365. https://doi.org/10.1177/1356389019852126

Bruno, P., & Strunk, K. O. (2019). Making the Cut: The Effectiveness of Teacher Screening and Hiring in the Los Angeles Unified School District. *Educational Evaluation and Policy Analysis*, *41*(4), 426–460. https://doi.org/10.3102/0162373719865561

Buerger, S., Kroehne, U., Koehler, C., & Goldhammer, F. (2019). What makes the difference? The impact of item properties on mode effects in reading assessments. *Studies in Educational Evaluation*, *62*(April), 1–9. https://doi.org/10.1016/j.stueduc.2019.04.005

Bührer, S., & Wroblewski, A. (2019). The practice and perceptions of RRI—A gender perspective. *Evaluation and Program Planning*, *77*(June), 101717. https://doi.org/10.1016/j.evalprogplan.2019.101717

Burd, M. B. (2019). Build and Improve Business Processes for Evaluation Consulting. *New Directions for Evaluation*, *2019*(164), 69–80. https://doi.org/10.1002/ev.20385

Buregeya, J. M., Loignon, C., & Brousselle, A. (2019). Contribution to healthy places: Risks of equity free health impact assessment. *Evaluation and Program Planning*, *73*(December 2018), 138–145. https://doi.org/10.1016/j.evalprogplan.2018.12.007

Burger, K. (2019). The socio-spatial dimension of educational inequality: A comparative European analysis. *Studies in Educational Evaluation*, *62*(March), 171–186. https://doi.org/10.1016/j.stueduc.2019.03.009

Buriak, S. E., & Ayars, C. L. (2019). Evaluation of a drug and alcohol safety education program in aviation using interrupted time series and the Kirkpatrick framework. *Evaluation and Program Planning*, *73*(February 2018), 62–70. https://doi.org/10.1016/j.evalprogplan.2018.11.003

Burke, H. M. C., Field, S., González-Calvo, L., Eichleay, M. A., & Moon, T. D. (2019). Quasi-experimental evaluation using confirmatory procedures: A case study of an economic and social empowerment intervention to reduce girls' vulnerability to HIV in rural Mozambique. *Evaluation and Program Planning*, *77*(December 2018), 101721. https://doi.org/10.1016/j.evalprogplan.2019.101721

Burke, H. M. C., Packer, C., González-Calvo, L., Ridgeway, K., Lenzi, R., Green, A. F., & Moon, T. D. (2019). A longitudinal qualitative evaluation of an economic and social empowerment intervention to reduce girls' vulnerability to HIV in rural Mozambique. *Evaluation and Program Planning*, *77*(December 2018), 101682. https://doi.org/10.1016/j.evalprogplan.2019.101682

Caldwell, L. D., & Bledsoe, K. L. (2019). Can Social Justice Live in a House of Structural Racism? A Question for the Field of Evaluation. *American Journal of Evaluation*, *40*(1), 6–18. https://doi.org/10.1177/1098214018815772

Camilli, S. (2019). Big Shoes to Fill: An Evaluation Journey in the Footsteps of Daniel L. Stufflebeam. *Journal of MultiDisciplinary Evaluation*, *15*(33), 30–36. Retrieved from https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/609/489

Canham, S. L., Fang, M. L., Battersby, L., & Wada, M. (2019). Understanding the functionality of housing-related support services through mapping methods and dialogue. *Evaluation and Program Planning*, *72*(September 2018), 33–39. https://doi.org/10.1016/j.evalprogplan.2018.09.008

Canham, S. L., Wister, A., & O'Dea, E. (2019). Strengths, weaknesses, opportunities, and threats to housing first in Metro Vancouver. *Evaluation and Program Planning*, *75*(May), 69–77. https://doi.org/10.1016/j.evalprogplan.2019.05.004

Cao, Z., Yu, S., & Huang, J. (2019). A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study. *Studies in Educational Evaluation*, *63*(September 2018), 102–112. https://doi.org/10.1016/j.stueduc.2019.08.001

Chachu, D. O. (2019). Implementation of a project-based comprehensive monitoring and evaluation strategy in Ghana and Côte d'Ivoire: Key lessons. *African Evaluation Journal*, *7*(1), 1–7. https://doi.org/10.4102/aej.v7i1.387

Chandna, K., Vine, M. M., Snelling, S., Harris, R., Smylie, J., & Manson, H. (2019). Principles, Approaches, and Methods for Evaluation in Indigenous Contexts: A Grey Literature Scoping Review. *Canadian Journal of Program Evaluation*, *34*(1), 21–47. https://doi.org/10.3138/cjpe.43050

Chang, C. C., Su, J. A., Chang, K. C., Lin, C. Y., Koschorke, M., Rüsch, N., & Thornicroft, G. (2019). Development of the Family Stigma Stress Scale (FSSS) for Detecting Stigma Stress in Caregivers of People With Mental Illness. *Evaluation and the Health Professions*, *42*(2), 148–168. https://doi.org/10.1177/0163278717745658

Chechak, D. J., Dunlop, J. M., & Holosko, M. J. (2019). Evaluating Youth Drop-In Programs: The Utility of Process Evaluation Methods. *Canadian Journal of Program Evaluation*, *34*(1), 152–164. https://doi.org/10.3138/cjpe.42976

Chelimsky, E. (2019). Improving the Match Between Sustainability Questions and Evaluation Practice: Some Reflections and a Checklist. *New Directions for Evaluation*, *2019*(162), 69–86. https://doi.org/10.1002/ev.20363

Chen, H. T., Morosanu, L., Powell-Threets, K., Lian, B., & Turner, N. (2019). Assessment of the monitoring and evaluation system of a population-based program: Theory-driven evaluation approach. *Evaluation and Program Planning*, *77*(August), 101719. https://doi.org/10.1016/j.evalprogplan.2019.101719

Chen, H., Su, Y., Zheng, Q., & Li, L. (2019). Towards an operationalized and effective school self-evaluation system: An ongoing action research study in Chinese secondary schools since 2011. *Studies in Educational Evaluation*, *60*(August 2018), 117–129. https://doi.org/10.1016/j.stueduc.2018.12.005

Cheng, M. W. T., & Chan, C. K. Y. (2019). An experimental test: Using rubrics for reflective writing to develop reflection. *Studies in Educational Evaluation*, *61*(April), 176–182. https://doi.org/10.1016/j.stueduc.2019.04.001

Chou, C. F., Vallin, T., Rogers, S., Streeter, R. A., Worede, I., Washko, M. M., … Ungard, J. (2019). Impact of the Health Resources and Services Administration's state oral health workforce grant program on dental workforce density and access to oral health care. *Evaluation and Program Planning*, *75*(September 2018), 43–53. https://doi.org/10.1016/j.evalprogplan.2019.04.004

Christie, C. A., & Alkin, M. C. (2019). Editors' Notes. *New Directions for Evaluation*, *2019*(163), 7–10. https://doi.org/10.1002/ev.20371

Christie, C. A., & Lemire, S. T. (2019). Why Evaluation Theory Should Be Used to Inform Evaluation Policy. *American Journal of Evaluation*, *40*(4), 490–508. https://doi.org/10.1177/1098214018824045

Cimpian, J. R., Marsh, J. A., Martorell, P., & Polikoff, M. S. (2019). Editors' Introduction. *Educational Evaluation and Policy Analysis*, *41*(1), 3–4. https://doi.org/10.3102/0162373718806855

Clark, N. (2019). Building a Culturally Responsive Independent Consulting Practice. *New Directions for Evaluation*, *2019*(164), 81–88. https://doi.org/10.1002/ev.20382

Cloete, F., & Auriacombe, C. (2019). Revisiting decoloniality for more effective research and evaluation. *African Evaluation Journal*, *7*(1), 1–10. https://doi.org/10.4102/aej.v7i1.363

Coldwell, M. (2019). Reconsidering context: Six underlying features of context to improve learning from evaluation. *Evaluation*, *25*(1), 99–117. https://doi.org/10.1177/1356389018803234

Coll, K. M., Sawyer, S., Scholl, S., & Hauser, N. (2019). A logic model development for an adolescent based intervention to improve benefits from Therapeutic Residential Care (TRC). *Evaluation and Program Planning*, *76*(June), 101678. https://doi.org/10.1016/j.evalprogplan.2019.101678

Conway, A., O'Donnell, C., & Yates, P. (2019). *The Effectiveness of the Nurse Care Coordinator Role on Patient-Reported and Health Service Outcomes: A Systematic Review*. *Evaluation and the Health Professions* (Vol. 42). https://doi.org/10.1177/0163278717734610

Cook, C., Camman, C., Leyland, A., O'Shea, S., & Towle, A. (2019). Dramatizing learning, performing ourselves: Stories of theatre-based evaluation in Vancouver's downtown eastside. *Canadian Journal of Program Evaluation*, *34*(2), 255–271. https://doi.org/10.3138/cjpe.44172

Cortes, K. E., & Lincove, J. A. (2019). Match or Mismatch? Automatic Admissions and College Preferences of Low- and High-Income Students. *Educational Evaluation and Policy Analysis*, *41*(1), 98–123. https://doi.org/10.3102/0162373718813360

Cox, K., & Kelcey, B. (2019). Optimal Design of Cluster- and Multisite-Randomized Studies Using Fallible Outcome Measures. *Evaluation Review*, *43*(3–4), 189–225. https://doi.org/10.1177/0193841X19870878

Cuenca-García, E., Sánchez, A., & Navarro-Pabsdorf, M. (2019). Assessing the performance of the least developed countries in terms of the Millennium Development Goals. *Evaluation and Program Planning*, *72*(September 2018), 54–66. https://doi.org/10.1016/j.evalprogplan.2018.09.009

Curran, L., & Taylor-Barnett, P. (2019). Evaluating projects in multifaceted and marginalised communities: The need for mixed approaches. *Evaluation Journal of Australasia*, *19*(1), 22–38. https://doi.org/10.1177/1035719X19832688

Curry, D. W. (2019). Perspectives on Monitoring and Evaluation. *American Journal of Evaluation*, *40*(1), 147–150. https://doi.org/10.1177/1098214018775845

Czechowski, K., Sylvestre, J., & Moreau, K. (2019). Secure Data Handling: An Essential Competence for Evaluators. *Canadian Journal of Program Evaluation*, *1*, 139–151. https://doi.org/10.3138/cjpe.43319

Daas, R., Dijkstra, A. B., & Karsten, S. (2019). Assessing young people's citizenship attitudes using rubrics. *Studies in Educational Evaluation*, *62*(June), 118–128. https://doi.org/10.1016/j.stueduc.2019.05.002

Dahler-Larsen, P., & Boodhoo, A. (2019). Evaluation culture and good governance: Is there a link? *Evaluation*, *25*(3), 277–293. https://doi.org/10.1177/1356389018819110

Dare, L., & Nowicki, E. (2019). Engaging children and youth in research and evaluation using group concept mapping. *Evaluation and Program Planning*, *76*(November 2018), 101680. https://doi.org/10.1016/j.evalprogplan.2019.101680

Das, B. M., & Sartore-Baldwin, M. L. (2019). Development of a logic model for a service learning, dog walking course for college students. *Evaluation and Program Planning*, *76*(May), 101667. https://doi.org/10.1016/j.evalprogplan.2019.05.002

Daumiller, M., Siegel, S., & Dresel, M. (2019). Construction and validation of a short multidisciplinary research performance questionnaire (SMRPQ). *Research Evaluation*, *28*(3), 241–252. https://doi.org/10.1093/reseval/rvz009

Davari Torshizi, M., & Bahraman, M. (2019). I explain, therefore I learn: Improving students' assessment literacy and deep learning by teaching. *Studies in Educational Evaluation*, *61*(February), 66–73. https://doi.org/10.1016/j.stueduc.2019.03.002

De Oliveira, T. M., Amaral, L., & Pacheco, R. C. D. S. (2019). Multi/inter/transdisciplinary assessment: A systemic framework proposal to evaluate graduate courses and research teams. *Research Evaluation*, *28*(1), 23–36. https://doi.org/10.1093/reseval/rvy013

Dean-Coffey, J. (2019). Book review. *Evaluation and Program Planning*, *75*, 89–90. https://doi.org/10.1016/j.evalprogplan.2019.04.001

Dee, T. S., & Dizon-Ross, E. (2019). *School Performance, Accountability, and Waiver Reforms: Evidence From Louisiana*. *Educational Evaluation and Policy Analysis* (Vol. 41). https://doi.org/10.3102/0162373719849944

Degn, L., Mejlgaard, N., & Schneider, J. W. (2019). Using mixed methods to map vaguely defined research areas. *Research Evaluation*, *28*(4), 394–404. https://doi.org/10.1093/reseval/rvz025

Delahais, T., & Lacouette-Fougère, C. (2019). Try again. Fail again. Fail better. Analysis of the contribution of 65 evaluations to the modernisation of public action in France. *Evaluation*, *25*(2), 131–148. https://doi.org/10.1177/1356389018823237

DeLancey, D. (2019). Introduction to articles prepared by CES Calgary 2018 keynote panel members on reconciliation and culturally responsive evaluation-Rhetoric or reality? *Canadian Journal of Program Evaluation*, *34*(2), 329–330. https://doi.org/10.3138/cjpe.67979

DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019). Assessment mindset: Exploring the relationship between teacher mindset and approaches to classroom assessment. *Studies in Educational Evaluation*, *61*(March), 159–169. https://doi.org/10.1016/j.stueduc.2019.03.012

Deutschlander, D. (2019). Enhancing Engagement With Faculty and Staff to Facilitate Student Success: An Evaluation of a Parent Intervention. *Educational Evaluation and Policy Analysis*, *41*(3), 239–259. https://doi.org/10.3102/0162373719845653

Dillman, L., Scott, C., & Kinarsky, A. R. (2019). The Case Study: The Women Affirming Motherhood Program. *New Directions for Evaluation*, *2019*(163), 19–26. https://doi.org/10.1002/ev.20376

Dingyloudi, F., Strijbos, J. W., & de Laat, M. F. (2019). Value creation: What matters most in Communities of Learning Practice in higher education. *Studies in Educational Evaluation*, *62*(September 2018), 209–223. https://doi.org/10.1016/j.stueduc.2019.05.006

Dinh, K., Worth, H., & Haire, B. (2019). Buddhist evaluation: Applying a Buddhist world view to the Most Significant Change Technique. *Evaluation*, *25*(4), 477–495. https://doi.org/10.1177/1356389019841654

Dinh, K., Worth, H., Haire, B., & Hong, K. T. (2019). Confucian Evaluation: Reframing Contribution Analysis Using a Confucian Lens. *American Journal of Evaluation*, *40*(4), 562–574. https://doi.org/10.1177/1098214018813008

Donaldson, S. I. (2019). Where do we stand? Recent AEA member views on professionalization. *Evaluation and Program Planning*, *72*(September 2018), 152–161. https://doi.org/10.1016/j.evalprogplan.2018.10.009

Downes, A., Novicki, E., & Howard, J. (2019). Using the Contribution Analysis Approach to Evaluate Science Impact: A Case Study of the National Institute for Occupational Safety and Health. *American Journal of Evaluation*, *40*(2), 177–189. https://doi.org/10.1177/1098214018767046

DuBow, W. M., & Litzler, E. (2019). The Development and Use of a Theory of Change to Align Programs and Evaluation in a Complex, National Initiative. *American Journal of Evaluation*, *40*(2), 231–248. https://doi.org/10.1177/1098214018778132

Ebenso, B., Manzano, A., Uzochukwu, B., Etiaba, E., Huss, R., Ensor, T., … Mirzoev, T. (2019). Dealing with context in logic model development: Reflections from a realist evaluation of a community health worker programme in Nigeria. *Evaluation and Program Planning*, *73*(September 2018), 97–110. https://doi.org/10.1016/j.evalprogplan.2018.12.002

Eboreime, E. A., Idika, O., Omitiran, K., Eboreime, O., & Ibisomi, L. (2019). Primary healthcare planning, bottleneck analysis and performance improvement: An evaluation of processes and outcomes in a Nigerian context. *Evaluation and Program Planning*, *77*(April), 101712. https://doi.org/10.1016/j.evalprogplan.2019.101712

Ecalle, J., Gomes, C., Auphan, P., Cros, L., & Magnan, A. (2019). Effects of policy and educational interventions intended to reduce difficulties in literacy skills in grade 1. *Studies in Educational Evaluation*, *61*(February), 12–20. https://doi.org/10.1016/j.stueduc.2019.02.001

Eckhard, S., & Jankauskas, V. (2019). The politics of evaluation in international organizations: A comparative study of stakeholder influence potential. *Evaluation*, *25*(1), 62–79. https://doi.org/10.1177/1356389018803967

Editorial Board. (2019). Editorial. *Evaluation Journal of Australasia*, *19*(1), 3–7. https://doi.org/10.1177/1035719X19833775

Editorial Board. (2019). Publisher's Note. *Evaluation and Program Planning*, *73*, 232. https://doi.org/10.1016/j.evalprogplan.2019.02.007

Egan, C. A., Webster, C. A., Stewart, G. L., Weaver, R. G., Russ, L. B., Brian, A., & Stodden, D. F. (2019). Case study of a health optimizing physical education-based comprehensive school physical activity program. *Evaluation and Program Planning*, *72*(September 2018), 106–117. https://doi.org/10.1016/j.evalprogplan.2018.10.006

Eklund, M., & Brunt, D. (2019). Measuring Opportunities for Engaging in Meaningful Home-Based Activities in Housing Services for People With Psychiatric Disabilities: Development of the Perceived Meaning of Activity in Housing (PMA-H). *Evaluation and the Health Professions*, *42*(2), 131–147. https://doi.org/10.1177/0163278717727333

El Soufi, N., & See, B. H. (2019). Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Studies in Educational Evaluation*, *60*(August 2018), 140–162. https://doi.org/10.1016/j.stueduc.2018.12.006

El-Den, S., O'Reilly, C. L., & Chen, T. F. (2019). Development and Psychometric Evaluation of a Questionnaire to Measure Attitudes Toward Perinatal Depression and Acceptability of Screening: The PND Attitudes and Screening Acceptability Questionnaire (PASAQ). *Evaluation and the Health Professions*, *42*(4), 498–522. https://doi.org/10.1177/0163278718801434

Elvbakken, K. T., & Hansen, H. F. (2019). Evidence producing organizations: Organizational translation of travelling evaluation ideas. *Evaluation*, *25*(3), 261–276. https://doi.org/10.1177/1356389018803965

Esko, T., & Miettinen, R. (2019). Scholarly understanding, mediating artefacts and the social impact of research in the educational sciences. *Research Evaluation*, *28*(4), 295–303. https://doi.org/10.1093/reseval/rvz018

Esparza Masana, R., & Fernández, T. (2019). Monitoring S3: Key dimensions and implications. *Evaluation and Program Planning*, *77*(July 2018), 101720. https://doi.org/10.1016/j.evalprogplan.2019.101720

Evans, S. H., & Clarke, P. (2019). Resolving design issues in developing a nutrition app: A case study using formative research. *Evaluation and Program Planning*, *72*(September 2018), 97–105. https://doi.org/10.1016/j.evalprogplan.2018.10.010

Evergreen, S., & Sabarre, N. (2019). Branding for the Independent Consultant: Basic to Advanced. *New Directions for Evaluation*, *2019*(164), 101–113. https://doi.org/10.1002/ev.20392

Fabra-Mata, J., & Mygind, J. (2019). Big data in evaluation: Experiences from using Twitter analysis to evaluate Norway's contribution to the peace process in Colombia. *Evaluation*, *25*(1), 6–22. https://doi.org/10.1177/1356389018804259

Feinstein, O. (2019). The entanglement of facts and values: An example from an agricultural project in Honduras. *Evaluation Journal of Australasia*, *19*(4), 199–203. https://doi.org/10.1177/1035719X19893081

Feldmann, M. L. (2019). Standing on the Shoulders of Giants: Applying Business Management Approaches to Independent Evaluation Consulting. *New Directions for Evaluation*, *2019*(164), 55–67. https://doi.org/10.1002/ev.20383

Fierro, R. S. (2019). Authentic Self in Evaluation Consulting: Reflections after 10 Years. *New Directions for Evaluation*, *2019*(164), 141–154. https://doi.org/10.1002/ev.20379

Filomeno, F. A. (2019). The potential of dialogues on social identity and diversity for immigrant civic integration. *Evaluation and Program Planning*, *77*(April), 101723. https://doi.org/10.1016/j.evalprogplan.2019.101723

Fleury, M. J., Grenier, G., Bamvita, J. M., & Chiocchio, F. (2019). Variables Associated With Perceived Work Role Performance Among Professionals in Multidisciplinary Mental Health Teams Overall and in Primary Care and Specialized Service Teams, Respectively. *Evaluation and the Health Professions*, *42*(2), 169–195. https://doi.org/10.1177/0163278717734282

Foo, C. Y., Reidpath, D. D., & Sivasampu, S. (2019). The Association Between Hospital Characteristics and Nonresponse in an Organization Survey: An Analysis of the National Healthcare Establishment and Workforce Survey in Malaysia. *Evaluation and the Health Professions*, *42*(1), 3–23. https://doi.org/10.1177/0163278717713569

Ford, T. G. (2019). Can the use of informal control mechanisms increase trust among teachers? An evaluation of the Accelerated Schools intervention. *Studies in Educational Evaluation*, *63*(June), 59–71. https://doi.org/10.1016/j.stueduc.2019.06.001

Ford-Paz, R. E., Crown, L., Lawton, K., Goldenthal, H., Day, G., Coyne, C. A., … Cicchetti, C. (2019). Working on Womanhood (WOW): A participatory formative evaluation of a community-developed intervention. *Evaluation and Program Planning*, *72*(September 2018), 237–249. https://doi.org/10.1016/j.evalprogplan.2018.10.007

Freer, G., & Lemire, S. (2019). Can't see the wood for the logframe: Integrating logframes and theories of change in development evaluation. *Canadian Journal of Program Evaluation*, *33*(3), 336–353. https://doi.org/10.3138/cjpe.53007

Froncek, B., & Rohmann, A. (2019). "You Get the Great Feeling That You're Being Heard But in the End You Realize That Things Will Be Done Differently and in Others' Favor": An Experimental Investigation of Negative Effects of Participation in Evaluation. *American Journal of Evaluation*, *40*(1), 19–34. https://doi.org/10.1177/1098214018813447

Fu, A. C., Kannan, A., & Shavelson, R. J. (2019). Direct and Unobtrusive Measures of Informal STEM Education Outcomes. *New Directions for Evaluation*, *2019*(161), 35–57. https://doi.org/10.1002/ev.20348

Fu, A. C., Kannan, A., & Shavelson, R. J. (2019). Editors' Notes. *New Directions for Evaluation*, *2019*(161), 7–15. https://doi.org/10.1002/ev.20349

Fu, A. C., Kannan, A., & Shavelson, R. J. (2019). Synthesis of Issues and Future Directions. *New Directions for Evaluation*, *2019*(161), 125–131. https://doi.org/10.1002/ev.20350

Galster, G. C., Santiago, A. M., Smith, R. J., & Leroux, J. (2019). Benefit–Cost Analysis of an Innovative Program for Self-Sufficiency and Homeownership. *Evaluation Review*, *43*(1–2), 3–40. https://doi.org/10.1177/0193841X19846697

Galstyan, E., Galimov, A., & Sussman, S. (2019). Commentary: The Emergence of Pod Mods at Vape Shops. *Evaluation and the Health Professions*, *42*(1), 118–124. https://doi.org/10.1177/0163278718812976

Gao, R., Liu, J., Johnson, R., Wang, J., & Hu, L. (2019). Validating an ethical decision-making model of assessment using authentic scenarios. *Studies in Educational Evaluation*, *62*(May), 187–196. https://doi.org/10.1016/j.stueduc.2019.05.003

Gao, X., Shen, J., Wu, H., & Krenn, H. Y. (2019). Evaluating program effects: Conceptualizing and demonstrating a typology. *Evaluation and Program Planning*, *72*(September 2018), 88–96. https://doi.org/10.1016/j.evalprogplan.2018.10.008

García, T., Boom, J., Kroesbergen, E. H., Núñez, J. C., & Rodríguez, C. (2019). Planning, execution, and revision in mathematics problem solving: Does the order of the phases matter? *Studies in Educational Evaluation*, *61*(February), 83–93. https://doi.org/10.1016/j.stueduc.2019.03.001

Garibay, C., & Teasdale, R. M. (2019). Equity and Evaluation in Informal STEM Education. *New Directions for Evaluation*, *2019*(161), 87–106. https://doi.org/10.1002/ev.20352

Garney, W. R., Wilson, K., Nelon, J. L., Hays, C., Garcia, K. M., Muraleetharan, D., … McLeroy, K. (2019). Understanding innovation in health program planning and development. *Evaluation and Program Planning*, *73*(February), 226–231. https://doi.org/10.1016/j.evalprogplan.2019.02.001

Gavarkovs, A. G., Blunt, W., & Petrella, R. J. (2019). A protocol for designing online training to support the implementation of community-based interventions. *Evaluation and Program Planning*, *72*(August 2018), 77–87. https://doi.org/10.1016/j.evalprogplan.2018.10.013

Germuth, A. A. (2019). Succeeding as an Independent Evaluation Consultant: Requisite Skills and Attributes. *New Directions for Evaluation*, *2019*(164), 43–54. https://doi.org/10.1002/ev.20386

Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2019). Students' perceptions of assessment quality related to their learning approaches and learning outcomes. *Studies in Educational Evaluation*, *63*(August), 72–82. https://doi.org/10.1016/j.stueduc.2019.07.005

Gibson, C., Stutchbury, T., Ikutegbe, V., & Michielin, N. (2019). Challenge-led interdisciplinary research in practice: Program design, early career research, and a dialogic approach to building unlikely collaborations. *Research Evaluation*, *28*(1), 51–62. https://doi.org/10.1093/reseval/rvy039

Gingras, Y., & Khelfaoui, M. (2019). Do we need a book citation index for research evaluation? *Research Evaluation*, *28*(4), 383–393. https://doi.org/10.1093/reseval/rvz024

Glazer, J. L., Massell, D., & Malone, M. (2019). Charter Schools in Turnaround: Competing Institutional Logics in the Tennessee Achievement School District. *Educational Evaluation and Policy Analysis*, *41*(1), 5–33. https://doi.org/10.3102/0162373718795051

Glegg, S. M. N., Ryce, A., & Brownlee, K. (2019). A visual management tool for program planning, project management and evaluation in paediatric health care. *Evaluation and Program Planning*, *72*(January 2018), 16–23. https://doi.org/10.1016/j.evalprogplan.2018.09.005

Glennie, M., O'Donnell, M., Brown, M., & Benson, J. (2019). Referees or sponsors? the role of evaluators in the promotion of research scientists in a public research organization. *Research Evaluation*, *28*(1), 63–72. https://doi.org/10.1093/reseval/rvy035

Glock, S., & Kleen, H. (2019). Attitudes toward students from ethnic minority groups: The roles of preservice teachers' own ethnic backgrounds and teacher efficacy activation. *Studies in Educational Evaluation*, *62*(May), 82–91. https://doi.org/10.1016/j.stueduc.2019.04.010

Goldman, I., Deliwe, C. N., Taylor, S., Ishmail, Z., Smith, L., Masangu, T., … Robertsen, J. (2019). Evaluation2 – Evaluating the national evaluation system in South Africa: What has

been achieved in the first 5 years? *African Evaluation Journal*, *7*(1), 1–11. https://doi.org/10.4102/aej.v7i1.400

Gonot-Schoupinsky, F. N., & Garip, G. (2019). A flexible framework for planning and evaluating early-stage health interventions: FRAME-IT. *Evaluation and Program Planning*, *77*(March), 101685. https://doi.org/10.1016/j.evalprogplan.2019.101685

Gotch, C. M., & McLean, C. (2019). Teacher outcomes from a statewide initiative to build assessment literacy. *Studies in Educational Evaluation*, *62*(February), 30–36. https://doi.org/10.1016/j.stueduc.2019.04.003

Gould, L. (2019). Foreword: On 'values.' *Evaluation Journal of Australasia*, *19*(4), 157–158. https://doi.org/10.1177/1035719X19895150

Gould, L., Quadrelli, C., & Rossingh, B. (2019). Editorial. *Evaluation Journal of Australasia*, *19*(3), 113–114. https://doi.org/10.1177/1035719X19878910

Goyal, N., & Howlett, M. (2019). Combining internal and external evaluations within a multilevel evaluation framework: Computational text analysis of lessons from the Asian Development Bank. *Evaluation*, *25*(3), 366–380. https://doi.org/10.1177/1356389019827035

Grack Nelson, A., Goeke, M., Auster, R., Peterman, K., & Lussenhop, A. (2019). Shared Measures for Evaluating Common Outcomes of Informal STEM Education Experiences. *New Directions for Evaluation*, *2019*(161), 59–86. https://doi.org/10.1002/ev.20353

Grack Nelson, A., King, J. A., Lawrenz, F., Reich, C., Bequette, M., Pattison, S., … Francisco, M. (2019). Using a Complex Adaptive Systems Perspective to Illuminate the Concept of Evaluation Capacity Building in a Network. *American Journal of Evaluation*, *40*(2), 214–230. https://doi.org/10.1177/1098214018773877

Gray, J. S., Connolly, J. P., & Brown, M. A. (2019). Measuring intercultural knowledge and competence in college essays: Does a performance-based rubric have construct validity? *Studies in Educational Evaluation*, *62*(June), 142–148. https://doi.org/10.1016/j.stueduc.2019.05.007

Gremillion, H., & Powell, C. (2019). Evaluating Unitec's ALLY workshop on diverse sexuality and gender inclusivity. *Evaluation Journal of Australasia*, *19*(3), 134–146. https://doi.org/10.1177/1035719X19875595

Gullickson, A. M., & Hannum, K. M. (2019). Making values explicit in evaluation practice. *Evaluation Journal of Australasia*, *19*(4), 162–178. https://doi.org/10.1177/1035719X19893892

Gullickson, A. M., Hassall, K., Hannum, K. M., & Roorda, M. (2019). Editorial. *Evaluation Journal of Australasia*, *19*(4), 159–161. https://doi.org/10.1177/1035719X19894837

Gullickson, A. M., King, J. A., LaVelle, J. M., & Clinton, J. M. (2019). The current state of evaluator education: A situation analysis and call to action. *Evaluation and Program Planning*, *75*(February), 20–30. https://doi.org/10.1016/j.evalprogplan.2019.02.012

Guo, W., Lau, K. L., & Wei, J. (2019). Teacher feedback and students' self-regulated learning in mathematics: A comparison between a high-achieving and a low-achieving secondary schools. *Studies in Educational Evaluation*, *63*(May), 48–58. https://doi.org/10.1016/j.stueduc.2019.07.001

Ha, K. M. (2019). Integrating the resources of Korean disaster management research via the Johari window. *Evaluation and Program Planning*, *77*(June), 101724. https://doi.org/10.1016/j.evalprogplan.2019.101724

Hagood, L. P. (2019). The Financial Benefits and Burdens of Performance Funding in Higher Education. *Educational Evaluation and Policy Analysis*, *41*(2), 189–213. https://doi.org/10.3102/0162373719837318

Hakkola, L., Allan, E. J., & Kerschner, D. (2019). Applying utilization-focused evaluation to high school hazing prevention: A pilot intervention. *Evaluation and Program Planning*, *75*(May), 61–68. https://doi.org/10.1016/j.evalprogplan.2019.05.005

Hamzeh, J., Pluye, P., Bush, P. L., Ruchon, C., Vedel, I., & Hudon, C. (2019). Towards an assessment for organizational participatory research health partnerships: A systematic mixed studies review with framework synthesis. *Evaluation and Program Planning*, *73*(June 2018), 116–128. https://doi.org/10.1016/j.evalprogplan.2018.12.003

Hanna, F., Oostdam, R., Severiens, S. E., & Zijlstra, B. J. H. (2019). Primary student teachers' professional identity tensions: The construction and psychometric quality of the professional identity tensions scale. *Studies in Educational Evaluation*, *61*(February), 21–33. https://doi.org/10.1016/j.stueduc.2019.02.002

Hansson, S., & Polk, M. (2019). Comments to Belcher et al. 2018's critique of Hansson and Polk 2018. *Research Evaluation*, *28*(2), 202–205. https://doi.org/10.1093/reseval/rvz006

Harnar, M. (2019). Being an Evaluator: Your Practical Guide to Evaluation. *Journal of MultiDisciplinary Evaluation*, *15*(32), 55–58.

Harper, L. M., & Dickson, R. (2019). Using developmental evaluation principles to build capacity for knowledge mobilisation in health and social care. *Evaluation*, *25*(3), 330–348. https://doi.org/10.1177/1356389019840058

Harris, K., Henderson, S., & Wink, B. (2019). Mobilising Q methodology within a realist evaluation: Lessons from an empirical study. *Evaluation*, *25*(4), 430–448. https://doi.org/10.1177/1356389019841645

Hasford, J., Nelson, G., Worton, S. K., Macnaughton, E., MacLeod, T., Piat, M., … Goering, P. (2019). Knowledge translation and implementation of housing first in Canada: A qualitative assessment of capacity building needs for an evidence-based program. *Evaluation and Program Planning*, *75*(March), 1–9. https://doi.org/10.1016/j.evalprogplan.2019.03.001

Haynes, E., Marawili, M., Marika, B. M., Mitchell, A. G., Phillips, J., Bessarab, D., … Ralph, A. P. (2019). Community-based participatory action research on rheumatic heart disease in an Australian Aboriginal homeland: Evaluation of the 'On track watch' project. *Evaluation and Program Planning*, *74*(February), 38–53. https://doi.org/10.1016/j.evalprogplan.2019.02.010

Hecking, T., & Leydesdorff, L. (2019). Can topic models be used in research evaluations? Reproducibility, validity, and reliability when compared with semantic maps. *Research Evaluation*, *28*(3), 263–272. https://doi.org/10.1093/reseval/rvz015

Heinze, T., & Jappe, A. (2019). Evaluation of research in social sciences and humanities. *Research Evaluation*, *28*(2), 206–208. https://doi.org/10.1093/reseval/rvz005

Hendra, R., & Hill, A. (2019). Rethinking Response Rates: New Evidence of Little Relationship Between Survey Response Rates and Nonresponse Bias. *Evaluation Review*, *43*(5), 307–330. https://doi.org/10.1177/0193841X18807719

Heyeres, M., Tsey, K., Yang, Y., Yan, L., & Jiang, H. (2019). The characteristics and reporting quality of research impact case studies: A systematic review. *Evaluation and Program Planning*, *73*(March 2018), 10–23. https://doi.org/10.1016/j.evalprogplan.2018.11.002

Hickey, S., Couchman, K., Stapleton, H., Roe, Y., & Kildea, S. (2019). Experiences of health service providers establishing an Aboriginal-Mainstream partnership to improve maternity

care for Aboriginal and Torres Strait Islander families in an urban setting. *Evaluation and Program Planning*, *77*(March), 101705. https://doi.org/10.1016/j.evalprogplan.2019.101705

Hicks, D. (2019). Introduction to special section: Keynotes from STI 2017. *Research Evaluation*, *28*(1), 1. https://doi.org/10.1093/reseval/rvy031

Hilton, L. G., & Azzam, T. (2019). Crowdsourcing Qualitative Thematic Analysis. *American Journal of Evaluation*, *40*(4), 575–589. https://doi.org/10.1177/1098214019836674

Hinyard, L., Toomey, E., Eliot, K., & Breitbach, A. (2019). Student Perceptions of Collaboration Skills in an Interprofessional Context: Development and Initial Validation of the Self-Assessed Collaboration Skills Instrument. *Evaluation and the Health Professions*, *42*(4), 450–472. https://doi.org/10.1177/0163278717752438

Hiruy, K., Murphy, R., Lewis, T., White, W. T., & Purcell, S. W. (2019). Measuring scientific impact of fisheries and aquaculture research-for-development projects in South East Asia and the Pacific. *Research Evaluation*, *28*(4), 313–325. https://doi.org/10.1093/reseval/rvz019

Hobson, K. A., Coryn, C. L. S., Fierro, L. A., & Sherwood-Laughlin, C. M. (2019). Instruction of Evaluation Competencies in Council on Education for Public Health (CEPH)-Accredited Master of Public Health (MPH) Degree Programs. *American Journal of Evaluation*, *40*(4), 590–606. https://doi.org/10.1177/1098214019845510

Hohl, S. D., Knerr, S., & Thompson, B. (2019). A framework for coordination center responsibilities and performance in a multi-site, transdisciplinary public health research initiative. *Research Evaluation*, *28*(3), 279–289. https://doi.org/10.1093/reseval/rvz012

Hooton, C. A. (2019). The application of micro-geographic economic analysis in urban policy evaluation. *Evaluation and Program Planning*, *72*(August 2018), 125–135. https://doi.org/10.1016/j.evalprogplan.2018.10.002

Horta, H., Cattaneo, M., & Meoli, M. (2019). The impact of Ph.D. funding on time to Ph.D. completion. *Research Evaluation*, *28*(2), 182–195. https://doi.org/10.1093/reseval/rvz002

Hortigüela Arroyo, M., & Ubillos Landa, S. (2019). Evaluability assessment of a community development leisure program in Spain. *Evaluation and Program Planning*, *72*(September 2018), 219–226. https://doi.org/10.1016/j.evalprogplan.2018.10.014

House, E. R. (2019). Evaluation With a Focus on Justice. *New Directions for Evaluation*, *2019*(163), 61–72. https://doi.org/10.1002/ev.20378

Høydal, Ø. S. (2019). Evaluation influence: Method and societal phenomenon. *Evaluation*, *25*(3), 381–398. https://doi.org/10.1177/1356389018804263

Huebner, A., & Lucht, M. (2019). Generalizability theory in R. *Practical Assessment, Research and Evaluation*, *24*(5), 1–12.

Hugo, L., & Botma, Y. (2019). Looking beneath the surface of a preceptor-training programme through a realist evaluation. *Evaluation and Program Planning*, *73*(June 2018), 195–203. https://doi.org/10.1016/j.evalprogplan.2019.01.005

Hull, M., & Duch, K. (2019). One-to-One Technology and Student Outcomes: Evidence From Mooresville's Digital Conversion Initiative. *Educational Evaluation and Policy Analysis*, *41*(1), 79–97. https://doi.org/10.3102/0162373718799969

Hunt, E. T., Whitfield, M. L., Brazendale, K., Beets, M. W., & Weaver, R. G. (2019). Examining the impact of a summer learning program on children's weight status and cardiorespiratory fitness: A natural experiment. *Evaluation and Program Planning*, *74*(February), 84–90. https://doi.org/10.1016/j.evalprogplan.2019.02.009

Hussein, A., & Hussain, M. (2019). Social-media based assessment of academic programs. *Studies in Educational Evaluation*, *62*(June), 149–157. https://doi.org/10.1016/j.stueduc.2019.06.003

Ikin, K., & McClenaghan, P. (2019). Deep diving into values: A way to enhance empowerment evaluation. *Evaluation Journal of Australasia*, *19*(4), 204–210. https://doi.org/10.1177/1035719X19893057

Illahi, U., & Mir, M. S. (2019). Book review. *Evaluation and Program Planning*, *73*(January), 224–225. https://doi.org/10.1016/j.evalprogplan.2019.01.006

Ishola, F., & Cekan, J. (2019). Evaluating the sustainability of health programmes: A literature review. *African Evaluation Journal*, *7*(1), 1–7. https://doi.org/10.4102/aej.v7i1.369

Islam, M. S. (2019). An assessment of child protection in Bangladesh: How effective is NGO-led Child-Friendly Space? *Evaluation and Program Planning*, *72*(September 2018), 8–15. https://doi.org/10.1016/j.evalprogplan.2018.09.003

Ispas, S. A., & Iliescu, D. (2019). The Romanian Adaptation of the Survey Work–Home Interaction, NijmeGen. *Evaluation and the Health Professions*, *42*(2), 196–218. https://doi.org/10.1177/0163278718791303

Jabbar, H., Castro, A., & Germain, E. (2019). To Switch or Not to Switch? The Influence of School Choice and Labor Market Segmentation on Teachers' Job Searches. *Educational Evaluation and Policy Analysis*, *41*(3), 375–399. https://doi.org/10.3102/0162373719857689

Jarosewich, T., Feldmann, M. L., Martínez-Rubin, N., & Clark, N. (2019). Who We Are: Findings from the American Evaluation Association's Independent Consulting Topical Interest Group 2015 Decennial Survey. *New Directions for Evaluation*, *2019*(164), 27–41. https://doi.org/10.1002/ev.20387

Jescovitch, L. N., Scott, E. E., Cerchiara, J. A., Doherty, J. H., Wenderoth, M. P., Merrill, J. E., … Haudek, K. C. (2019). Deconstruction of holistic rubrics into analytic rubrics for large-scale assessments of students' reasoning of complex science concepts. *Practical Assessment, Research and Evaluation*, *24*(6). https://doi.org/10.7275/9h7f-mp76

Jiménez-Aguilar, A., Rodríguez-Oliveros, M. G., Uribe-Carvajal, R., González-Unzaga, M. A., Escalante-Izeta, E. I., & Reyes-Morales, H. (2019). Design of an educational strategy based on Intervention Mapping for nutritional health promotion in Child Care Centers. *Evaluation and Program Planning*, *76*(June), 101672. https://doi.org/10.1016/j.evalprogplan.2019.101672

Johnson, A. (2019). A Matter of Time: Variations in High School Course-Taking by Years-as-EL Subgroup. *Educational Evaluation and Policy Analysis*, *41*(4), 461–482. https://doi.org/10.3102/0162373719867087

Johnson, B., & Činčera, J. (2019). Development of the ecological concepts of energy flow and materials cycling in middle school students participating in earth education programs. *Studies in Educational Evaluation*, *63*(June), 94–101. https://doi.org/10.1016/j.stueduc.2019.08.003

Jonker, H., März, V., & Voogt, J. (2019). Collaboration in teacher design teams: Untangling the relationship between experiences of the collaboration process and perceptions of the redesigned curriculum. *Studies in Educational Evaluation*, *61*(March), 138–149. https://doi.org/10.1016/j.stueduc.2019.03.010

Juaneda-Ayensa, E., Olarte-Pascual, C., Clavel San Emeterio, M., & Pelegrín-Borondo, J. (2019). Developing new "Professionals": Service learning in marketing as an opportunity to

innovate in higher education. *Studies in Educational Evaluation*, *60*(January), 163–169. https://doi.org/10.1016/j.stueduc.2019.01.001

Julnes, G. (2019). Supporting Transitions to Sustainability: Evaluation for Managing Processes in the Public Interest. *New Directions for Evaluation*, *2019*(162), 119–154. https://doi.org/10.1002/ev.20366

Julnes, G. (2019). Editor's Notes. *New Directions for Evaluation*, *2019*(162), 7–11. https://doi.org/10.1002/ev.20359

Julnes, G. (2019). Evaluating Sustainability: Controversies, Challenges, and Opportunities. *New Directions for Evaluation*, *2019*(162), 13–28. https://doi.org/10.1002/ev.20361

Julnes, G. (2019). From the Editor. *American Journal of Evaluation*, *40*(2), 158–159. https://doi.org/10.1177/1098214019841576

Junge, N., & Negroustoueva, S. (2019). Bridging Divides and Creating Opportunities in International Evaluation Consulting. *New Directions for Evaluation*, *2019*(164), 127–139. https://doi.org/10.1002/ev.20388

Kallemeyn, L. M. (2019). Stakeholders' shifting values in an evaluation of a literacy professional development programme. *Evaluation Journal of Australasia*, *19*(4), 211–216. https://doi.org/10.1177/1035719X19893064

Karami, M., Pishghadam, R., & Baghaei, P. (2019). A probe into EFL learners' emotioncy as a source of test bias: Insights from differential item functioning analysis. *Studies in Educational Evaluation*, *60*(August 2018), 170–178. https://doi.org/10.1016/j.stueduc.2019.01.003

Karimi, M. N., & Norouzi, M. (2019). Developing and validating three measures of possible language teacher selves. *Studies in Educational Evaluation*, *62*(May), 49–60. https://doi.org/10.1016/j.stueduc.2019.04.006

Karjanto, N., & Simon, L. (2019). English-medium instruction Calculus in Confucian-Heritage Culture: Flipping the class or overriding the culture? *Studies in Educational Evaluation*, *63*(February), 122–135. https://doi.org/10.1016/j.stueduc.2019.07.002

Karras-Jean Gilles, J., Astuto, J., Gjicali, K., & Allen, L. R. (2019). Sample Retention in an Urban Context: Exploring Influential Factors Within a Longitudinal Randomized Evaluation. *American Journal of Evaluation*, *40*(2), 268–290. https://doi.org/10.1177/1098214017742719

Kastrinos, N. (2019). Our common history in the USA: a European perspective. *Research Evaluation*, *28*(3), 290–293. https://doi.org/10.1093/reseval/rvz004

Kelly, L. (2019). Children's opinion of retrospective pre-post 'then-test' survey validity. *Journal of MultiDisciplinary Evaluation*, 54–65.

Ken-Opurum, J., Lynch, K., Vandergraff, D., Miller, D. K., & Savaiano, D. A. (2019). A mixed-methods evaluation using effectiveness perception surveys, social network analysis, and county-level health statistics: A pilot study of eight rural Indiana community health coalitions. *Evaluation and Program Planning*, *77*(August), 101709. https://doi.org/10.1016/j.evalprogplan.2019.101709

Kim, S., Wallsworth, G., Xu, R., Schneider, B., Frank, K., Jacob, B., & Dynarski, S. (2019). The Impact of the Michigan Merit Curriculum on High School Math Course-Taking. *Educational Evaluation and Policy Analysis*, *41*(2), 164–188. https://doi.org/10.3102/0162373719834067

Kinarsky, A. R., Alkin, M. C., & Christie, C. A. (2019). What Have We Learned? *New Directions for Evaluation*, *2019*(163), 83–93. https://doi.org/10.1002/ev.20372

King, J. A., & Alkin, M. C. (2019). The Centrality of Use: Theories of Evaluation Use and Influence and Thoughts on the First 50 Years of Use Research. *American Journal of Evaluation*, *40*(3), 431–458. https://doi.org/10.1177/1098214018796328

King, J. (2019). Combining multiple approaches to valuing in the MUVA female economic empowerment programme. *Evaluation Journal of Australasia*, *19*(4), 217–225. https://doi.org/10.1177/1035719X19893080

Koleros, A., & Mayne, J. (2019). Using actor-based theories of change to conduct robust evaluation in complex settings. *Canadian Journal of Program Evaluation*, *33*(3), 292–315. https://doi.org/10.3138/cjpe.52946

Kooli, C. (2019). Governing and managing higher education institutions: The quality audit contributions. *Evaluation and Program Planning*, *77*(May), 101713. https://doi.org/10.1016/j.evalprogplan.2019.101713

Kosko, K. W. (2019). A multiplicative reasoning assessment for fourth and fifth grade students. *Studies in Educational Evaluation*, *60*(November 2017), 32–42. https://doi.org/10.1016/j.stueduc.2018.11.003

Krog Lind, J. (2019). The missing link: How university managers mediate the impact of a performance-based research funding system. *Research Evaluation*, *28*(1), 84–93. https://doi.org/10.1093/reseval/rvy038

Kubiak, S., Shamrova, D., & Comartin, E. (2019). Enhancing knowledge of adolescent mental health among law enforcement: Implementing youth-focused crisis intervention team training. *Evaluation and Program Planning*, *73*(February 2018), 44–52. https://doi.org/10.1016/j.evalprogplan.2018.11.006

Kulczycki, E. (2019). Field patterns of scientometric indicators use for presenting research portfolio for assessment. *Research Evaluation*, *28*(2), 169–181. https://doi.org/10.1093/reseval/rvy043

Lachance, L., Watson, C., Blais, D., Ungar, M., Healey, G., Salaffie, M., … Lagace, M. C. (2019). Strengthening child and youth programs: A look at inter-organizational mentoring strategies. *Evaluation and Program Planning*, *76*(June). https://doi.org/10.1016/j.evalprogplan.2019.101679

LaChenaye, J. M., Boyce, A. S., Van Draanen, J., & Everett, K. (2019). Community, Theory, and Guidance: Reflections on the Benefits of Peer-Mentoring in Evaluation. *Canadian Journal of Program Evaluation*, *34*(1), 102–117. https://doi.org/10.3138/cjpe.42118

Lacoe, J., & Steinberg, M. P. (2019). Do Suspensions Affect Student Outcomes? *Educational Evaluation and Policy Analysis*, *41*(1), 34–62. https://doi.org/10.3102/0162373718794897

Lafontaine, D., Dupont, V., Jaegers, D., & Schillings, P. (2019). Self-concept in reading: Factor structure, cross-cultural invariance and relationships with reading achievement in an international context (PIRLS 2011). *Studies in Educational Evaluation*, *60*(January 2018), 78–89. https://doi.org/10.1016/j.stueduc.2018.11.005

Lam, S., Dodd, W., Whynot, J., & Skinner, K. (2019). How is gender being addressed in the international development evaluation literature? A meta-evaluation. *Research Evaluation*, *28*(2), 158–168. https://doi.org/10.1093/reseval/rvy042

Lam, T. C. M., & Valencia, E. (2019). Retrospective Pretest and Counterfactual Self-Report: Different or Same ? *Journal of MultiDisciplinary Evaluation*, *15*(22), 37–53.

Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Educational Evaluation*, *61*(February), 112–122. https://doi.org/10.1016/j.stueduc.2019.03.003

LaVenia, K. N., & Burgoon, J. (2019). Urban community as resource: Evaluation of the mentors in Toledo schools program. *Evaluation and Program Planning*, *72*(September 2018), 118–124. https://doi.org/10.1016/j.evalprogplan.2018.10.011

Lee, I., Mak, P., & Yuan, R. E. (2019). Assessment as learning in primary writing classrooms: An exploratory study. *Studies in Educational Evaluation*, *62*(May), 72–81. https://doi.org/10.1016/j.stueduc.2019.04.012

Lee, Y. M., Song, J. E., Park, C., & Son, Y. J. (2019). Psychometric Evaluation of the Korean Version of Patient-Centered Care Scale for Hospital Nurses. *Evaluation and the Health Professions*, *42*(3), 344–365. https://doi.org/10.1177/0163278718805244

Lemire, S., & Christie, C. A. (2019). Meta-modeling: A theory-based synthesis approach. *Canadian Journal of Program Evaluation*, *33*(3), 395–413. https://doi.org/10.3138/cjpe.52945

Lemire, S., Whynot, J., & Montague, S. (2019). How we model matters: A manifesto for the next generation of program theorizing. *Canadian Journal of Program Evaluation*, *33*(3), 414–433. https://doi.org/10.3138/cjpe.53070

Lidegaard, M., Olsen, K. B., & Legg, S. J. (2019). How was a national moving and handling people guideline intended to work? The underlying programme theory. *Evaluation and Program Planning*, *73*(September 2018), 163–175. https://doi.org/10.1016/j.evalprogplan.2019.01.002

Lilleholt, L., Aaby, A., & Makransky, G. (2019). Students admitted to university based on a cognitive test and MMI are less stressed than students admitted based on GPA. *Studies in Educational Evaluation*, *61*(March), 170–175. https://doi.org/10.1016/j.stueduc.2019.03.005

Lin, L., Lam, J. W. I., & Tse, S. K. (2019). Test takers' strategy use and L2 Chinese reading test performance in mainland China: A structural equation approach. *Studies in Educational Evaluation*, *60*(January), 189–198. https://doi.org/10.1016/j.stueduc.2019.01.002

Lin, P. Y., & Lin, Y. C. (2019). Understanding how teachers practise inclusive classroom assessment. *Studies in Educational Evaluation*, *63*(August), 113–121. https://doi.org/10.1016/j.stueduc.2019.08.002

Litwok, D., & Peck, L. R. (2019). Variance Estimation in Evaluations With No-Shows: A Comparison of Methods. *American Journal of Evaluation*, *40*(1), 104–118. https://doi.org/10.1177/1098214017749318

Liu, L., Du, X., Zhang, Z., & Zhou, J. (2019). Effect of problem-based learning in pharmacology education: A meta-analysis. *Studies in Educational Evaluation*, *60*(November 2018), 43–58. https://doi.org/10.1016/j.stueduc.2018.11.004

Liu, Y., Visone, J., Mongillo, M. B., & Lisi, P. (2019). What matters to teachers if evaluation is meant to help them improve? *Studies in Educational Evaluation*, *61*(March), 41–54. https://doi.org/10.1016/j.stueduc.2019.01.006

Logue, A. W., Douglas, D., & Watanabe-Rose, M. (2019). Corequisite Mathematics Remediation: Results Over Time and in Different Contexts. *Educational Evaluation and Policy Analysis*, *41*(3), 294–315. https://doi.org/10.3102/0162373719848777

López, Á. R., Souto, J. E., & Noblejas, M. L. A. (2019). Improving teaching capacity to increase student achievement: The key role of communication competences in Higher Education. *Studies in Educational Evaluation*, *60*(April 2018), 205–213. https://doi.org/10.1016/j.stueduc.2018.10.002

Luo, J., Ordóñez-Matamoros, G., & Kuhlmann, S. (2019). The balancing role of evaluation mechanisms in organizational governance - The case of publicly funded research institutions. *Research Evaluation*, *28*(4), 344–354. https://doi.org/10.1093/reseval/rvz022

Luyten, H., & Bazo, M. (2019). Transformational leadership, professional learning communities, teacher learning and learner centred teaching practices; Evidence on their interrelations in Mozambican primary education. *Studies in Educational Evaluation*, *60*(November 2017), 14–31. https://doi.org/10.1016/j.stueduc.2018.11.002

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). *Strengthening the Research Base That Informs STEM Instructional Improvement Efforts: A Meta-Analysis*. *Educational Evaluation and Policy Analysis* (Vol. 41). https://doi.org/10.3102/0162373719849044

Ma, L. (2019). Money, morale, and motivation: A study of the Output-Based Research Support Scheme in University College Dublin. *Research Evaluation*, *28*(4), 304–312. https://doi.org/10.1093/reseval/rvz017

Ma, L., & Ladisch, M. (2019). Evaluation complacency or evaluation inertia? A study of evaluative metrics and research practices in Irish universities. *Research Evaluation*, *28*(3), 209–217. https://doi.org/10.1093/reseval/rvz008

Maack, S. C. (2019). Taking Down Your Shingle? Or Not …. *New Directions for Evaluation*, *2019*(164), 169–183. https://doi.org/10.1002/ev.20380

Maciel, R. R. B. T., Chiavegato, L. D., Marin, L. S., Portella, D. D. A., de Souza, M. C., Camelier, F. W. R., & Padula, R. S. (2019). Development of an e-health education program at the workplace using formative research – Technologies for improving quality of life. *Evaluation and Program Planning*, *73*(December 2018), 129–137. https://doi.org/10.1016/j.evalprogplan.2018.12.009

Madhanagopal, J., Singh, O. P., Mohan, V., Sathasivam, K. V., Omar, A. H., & Abdul Kadir, M. R. (2019). New Device for Intrinsic Hand Muscle Strength Measurement: An Alternative to Strain Gauge Handheld Dynamometer. *Evaluation and the Health Professions*, *42*(1), 103–113. https://doi.org/10.1177/0163278717727568

Makiva, M., Ile, I., & Fagbadebo, O. M. (2019). Evaluating transformation progress of historically disadvantaged South Africans: Programme perspective on the downstream petroleum industry. *African Evaluation Journal*, *7*(1), 1–12. https://doi.org/10.4102/aej.v7i1.373

Maloney, J. (2019). Book review: Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches. *Evaluation Journal of Australasia*, *19*(2), 107–109. https://doi.org/10.1177/1035719x19848528

Mapitsa, C. B., Khumalo, L., Engel, H., & Wooldridge, D. (2019). Can massive open online courses fill African evaluation capacity gaps? *African Evaluation Journal*, *7*(1), 1–7. https://doi.org/10.4102/aej.v7i1.372

Martinaitis, Ž., Christenko, A., & Kraučiūnienė, L. (2019). Evaluation systems: How do they frame, generate and use evidence? *Evaluation*, *25*(1), 46–61. https://doi.org/10.1177/1356389018802135

Martínez-Rubin, N. (2019). Independent Evaluation Consulting: An Evolving Field. *New Directions for Evaluation*, *2019*(164), 13–25. https://doi.org/10.1002/ev.20390

Martínez-Rubin, N., Germuth, A. A., & Feldmann, M. L. (2019). Editors' Notes. *New Directions for Evaluation*, *2019*(164), 9–12. https://doi.org/10.1002/ev.20389

Mascareño, M. (2019). Constellations of risk factors as moderators of the impact of a randomized intervention on students' reading skills in rural India. *Studies in Educational Evaluation*, *60*(June 2018), 66–77. https://doi.org/10.1016/j.stueduc.2018.11.007

Mason, S., & Azzam, T. (2019). In Need of an Attitude Adjustment? The Role of Data Visualization in Attitude Change and Evaluation Influence. *American Journal of Evaluation*, *40*(2), 249–267. https://doi.org/10.1177/1098214018778808

Mason, S., & Hunt, A. (2019). So What Do You Do? Exploring Evaluator Descriptions of Their Work. *American Journal of Evaluation*, *40*(3), 395–413. https://doi.org/10.1177/1098214018767049

Mayne, J. (2019). Revisiting contribution analysis. *Canadian Journal of Program Evaluation*, *34*(2), 171–191. https://doi.org/10.3138/cjpe.68004

Mbava, N. P., & Dahler-Larsen, P. (2019). Evaluation in African contexts: The promises of participatory approaches in theory-based evaluations. *African Evaluation Journal*, *7*(1), 1–9. https://doi.org/10.4102/aej.v7i1.383

McAlindon, K., Neal, J. W., Neal, Z. P., Mills, K. J., & Lawlor, J. (2019). The BOND Framework: A Practical Application of Visual Communication Design and Marketing to Advance Evaluation Reporting. *American Journal of Evaluation*, *40*(2), 291–305. https://doi.org/10.1177/1098214018771219

McCausland, R. (2019). 'I'm sorry but I can't take a photo of someone's capacity being built': Reflections on evaluation of Indigenous policy and programmes. *Evaluation Journal of Australasia*, *19*(2), 64–78. https://doi.org/10.1177/1035719X19848529

McConnell, J. (2019). Adoption for adaptation: A theory-based approach for monitoring a complex policy initiative. *Evaluation and Program Planning*, *73*(August 2018), 214–223. https://doi.org/10.1016/j.evalprogplan.2019.01.008

McCool-Myers, M. (2019). Implementing condom distribution programs in the United States: Qualitative insights from program planners. *Evaluation and Program Planning*, *74*(February), 20–26. https://doi.org/10.1016/j.evalprogplan.2019.02.006

McDonald, D., & Stirling, R. (2019). Evaluating the capacity building roles of the state and territory peak bodies in the Australian alcohol and other drug sector. *Evaluation Journal of Australasia*, *19*(1), 39–48. https://doi.org/10.1177/1035719X18823569

McGee, J. B., & Winters, M. A. (2019). Rethinking the Structure of Teacher Retirement Benefits: Analyzing the Preferences of Entering Teachers. *Educational Evaluation and Policy Analysis*, *41*(1), 63–78. https://doi.org/10.3102/0162373718798488

McKegg, K. (2019). White privilege and the decolonization work needed in evaluation to support indigenous sovereignty and self-determination. *Canadian Journal of Program Evaluation*, *34*(2), 357–367. https://doi.org/10.3138/cjpe.67978

McLean, R. K. D., & Sen, K. (2019). Making a difference in the real world? A meta-analysis of the quality of use-oriented research using the Research Quality Plus approach. *Research Evaluation*, *28*(2), 123–135. https://doi.org/10.1093/reseval/rvy026

McNeish, R., Rigg, K. K., Tran, Q., & Hodges, S. (2019). Community-based behavioral health interventions: Developing strong community partnerships. *Evaluation and Program Planning*, *73*(September 2018), 111–115. https://doi.org/10.1016/j.evalprogplan.2018.12.005

Meng, A., Borg, V., & Clausen, T. (2019). Enhancing the social capital in industrial workplaces: Developing workplace interventions using intervention mapping. *Evaluation and Program Planning*, *72*(September 2018), 227–236. https://doi.org/10.1016/j.evalprogplan.2018.11.007

Mérida Rios, L., Márquez Serrano, M., Jiménez Aguilar, A., Barboza Chacón, L., Rueda Neria, C. M., & Arenas Monreal, L. (2019). Promoting fruit, vegetable and simple water consumption among mothers and teachers of preschool children: An Intervention Mapping initiative. *Evaluation and Program Planning*, *76*(655), 101675. https://doi.org/10.1016/j.evalprogplan.2019.101675

Meyers, D. C., Domitrovich, C. E., Dissi, R., Trejo, J., & Greenberg, M. T. (2019). Supporting systemic social and emotional learning with a schoolwide implementation model. *Evaluation and Program Planning*, *73*(April 2018), 53–61. https://doi.org/10.1016/j.evalprogplan.2018.11.005

Miller, J. J., Benner, K., Donohue-Dioh, J., & Segress, M. (2019). Supporting collegiate foster youth and alumni: A mixed-method planning approach for higher education. *Evaluation and Program Planning*, *72*(September 2018), 67–76. https://doi.org/10.1016/j.evalprogplan.2018.10.005

Miller, R., Wake, D., Whittingham, J., & Adelson, J. L. (2019). The content and construct development of the PACT survey. *Studies in Educational Evaluation*, *61*(February), 1–11. https://doi.org/10.1016/j.stueduc.2019.01.007

Milzow, K., Reinhardt, A., Söderberg, S., & Zinöcker, K. (2019). Understanding the use and usability of research evaluation studies. *Research Evaluation*, *28*(1), 94–107. https://doi.org/10.1093/reseval/rvy040

Mingo, M. V., Goldberg, J., Castro, M. de los A., Fillol, M. P., Mongillo, M., & Bedregal, P. (2019). Supporting the caregiver-child dyad's relationship: An evaluation of implementation quality in the Chilean Crecer Jugando program. *Evaluation and Program Planning*, *76*(May), 101668. https://doi.org/10.1016/j.evalprogplan.2019.05.003

Modestino, A. S., & Paulsen, R. J. (2019). Reducing inequality summer by summer: Lessons from an evaluation of the Boston Summer Youth Employment Program. *Evaluation and Program Planning*, *72*(September 2018), 40–53. https://doi.org/10.1016/j.evalprogplan.2018.09.006

Mohammad, T., Azman, A., & Anderstone, B. (2019). The global three: A Malaysian lens on the challenges and opportunities facing restorative justice planning and implementation. *Evaluation and Program Planning*, *72*(September 2018), 1–7. https://doi.org/10.1016/j.evalprogplan.2018.09.007

Montague, S. (2019). Does your implementation fit your theory of change? *Canadian Journal of Program Evaluation*, *33*(3), 316–335. https://doi.org/10.3138/cjpe.53008

Montrosse-Moorhead, B. (2019). Book Review: Educational Goods: Values, Evidence, and Decision-Making. *American Journal of Evaluation*, *40*(4), 607–612. https://doi.org/10.1177/1098214019868303

Moore, G. F., Evans, R. E., Hawkins, J., Littlecott, H., Melendez-Torres, G. J., Bonell, C., & Murphy, S. (2019). From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation*, *25*(1), 23–45. https://doi.org/10.1177/1356389018803219

Moore, M., Parsons, B., & Jessup, P. (2019). How Evaluators Can Use a Complex Systems Lens to Get "Untrapped" From Limiting Beliefs and Assumptions. *American Journal of Evaluation*, *40*(1), 75–91. https://doi.org/10.1177/1098214018756578

Moore, T., Bakken, L. L., & Wallace, B. (2019). The importance of systems thinking, context, and inclusion when studying the feasibility of expanding the FoodShare (SNAP) outreach program to rural communities of Northern Wisconsin. *Evaluation and Program Planning*, *74*(February), 10–17. https://doi.org/10.1016/j.evalprogplan.2019.02.003

Morell, J. A. (2019). Linking Management and Evaluation: Project Schedules as Program Models. *American Journal of Evaluation*, *40*(4), 472–489. https://doi.org/10.1177/1098214018784198

Moses, L. G. (2019). Growing and Nurturing Your Evaluation Business When Your Independent Consultancy Becomes an Employer. *New Directions for Evaluation*, *2019*(164), 115–126. https://doi.org/10.1002/ev.20381

Muiru, O. W. (2019). The B movement in East Africa: A shift in the culture of business. *African Evaluation Journal*, *7*(1), 1–12. https://doi.org/10.4102/aej.v7i1.333

Mundia, L., & Metussin, H. (2019). Exploring factors that improve mathematics achievement in Brunei. *Studies in Educational Evaluation*, *60*(January 2018), 214–222. https://doi.org/10.1016/j.stueduc.2018.10.003

Nicolaou, M., & Atkinson, M. (2019). Do student and survey characteristics affect the quality of UK undergraduate medical education course evaluation? A systematic review of the literature. *Studies in Educational Evaluation*, *62*(May), 92–103. https://doi.org/10.1016/j.stueduc.2019.04.011

Nielsen, J. V., Bredahl, T. V. G., Bugge, A., Klakk, H., & Skovgaard, T. (2019). Implementation of a successful long-term school based physical education intervention: Exploring provider and programme characteristics. *Evaluation and Program Planning*, *76*(June), 101674. https://doi.org/10.1016/j.evalprogplan.2019.101674

Nielsen, T., & Dammeyer, J. (2019). Measuring higher education students' perceived stress: An IRT-based construct validity study of the PSS-10. *Studies in Educational Evaluation*, *63*(June), 17–25. https://doi.org/10.1016/j.stueduc.2019.06.007

Noel, H. J., & Huang, A. R. (2019). The Effect of Varying Incentive Amounts on Physician Survey Response. *Evaluation and the Health Professions*, *42*(1), 71–81. https://doi.org/10.1177/0163278718809844

Nordesjö, K. (2019). Made in Sweden: The translation of a European evaluation approach. *Evaluation*, *25*(2), 189–206. https://doi.org/10.1177/1356389018803220

Nordhagen, S., Nielsen, J., van Mourik, T., Smith, E., & Klemm, R. (2019). Fostering CHANGE: Lessons from implementing a multi-country, multi-sector nutrition-sensitive agriculture project. *Evaluation and Program Planning*, *77*(March), 101695. https://doi.org/10.1016/j.evalprogplan.2019.101695

Nordström, T., Andersson, U. B., Fälth, L., & Gustafson, S. (2019). Teacher inquiry of using assessments and recommendations in teaching early reading. *Studies in Educational Evaluation*, *63*(June), 9–16. https://doi.org/10.1016/j.stueduc.2019.06.006

Norwood, M. F., Lakhani, A., Maujean, A., Downes, M., Fullagar, S., McIntyre, M., … Kendall, E. (2019). Assessing emotional and social health using photographs: An innovative research method for rural studies and its applicability in a care-farming program for youth. *Evaluation and Program Planning*, *77*(March), 101707. https://doi.org/10.1016/j.evalprogplan.2019.101707

O'Brien, J., Thompson, M. S., & Hagler, D. (2019). *Using Generalizability Theory to Inform Optimal Design for a Nursing Performance Assessment*. *Evaluation and the Health Professions* (Vol. 42). https://doi.org/10.1177/0163278717735565

O'Brien, S., McNamara, G., O'Hara, J., & Brown, M. (2019). Irish teachers, starting on a journey of data use for school self-evaluation. *Studies in Educational Evaluation*, *60*(January 2018), 1–13. https://doi.org/10.1016/j.stueduc.2018.11.001

Onken, J., Aragon, R., & Calcagno, A. M. (2019). Geographically-related outcomes of U.S. funding for small business research and development: Results of the research grant programs of a component of the National Institutes of Health. *Evaluation and Program Planning*, *77*(January), 101696. https://doi.org/10.1016/j.evalprogplan.2019.101696

Onken, J., Miklos, A. C., Dorsey, T. F., Aragon, R., & Calcagno, A. M. (2019). Using database linkages to measure innovation, commercialization, and survival of small businesses. *Evaluation and Program Planning*, *77*(August), 101710. https://doi.org/10.1016/j.evalprogplan.2019.101710

Onnis, L. ann, Hakendorf, M., Diamond, M., & Tsey, K. (2019). CQI approaches for evaluating management development programs: A case study with health service managers from geographically remote settings. *Evaluation and Program Planning*, *74*(February), 91–101. https://doi.org/10.1016/j.evalprogplan.2019.03.003

Ouda, J. B., Mulaudzi, P., Najoli, E. K., Wanyama, R., & Runhare, T. (2019). An evaluation of stakeholder capacity in the implementation of millenium village primary school meal project. *Evaluation and Program Planning*, *72*(July 2018), 179–187. https://doi.org/10.1016/j.evalprogplan.2018.10.003

Owen, J., & Rossingh, B. (2019). Evaluator perspective. *Evaluation Journal of Australasia*, *19*(3), 147–150. https://doi.org/10.1177/1035719X19874631

Owusu-Addo, E., Renzaho, A. M. N., & Smith, B. J. (2019). Cash transfers and the social determinants of health: Towards an initial realist program theory. *Evaluation*, *25*(2), 224–244. https://doi.org/10.1177/1356389018814868

Ozeki, S., Coryn, C. L. S., & Schröter, D. C. (2019). Evaluation logic in practice: Findings from two empirical investigations of American Evaluation Association members. *Evaluation and Program Planning*, *76*(June), 101681. https://doi.org/10.1016/j.evalprogplan.2019.101681

Palaco, I., Park, M. J., Kim, S. K., & Rho, J. J. (2019). Public–private partnerships for e-government in developing countries: An early stage assessment framework. *Evaluation and Program Planning*, *72*(September 2018), 205–218. https://doi.org/10.1016/j.evalprogplan.2018.10.015

Palmén, R., & Kalpazidou Schmidt, E. (2019). Analysing facilitating and hindering factors for implementing gender equality interventions in R&I: Structures and processes. *Evaluation and Program Planning*, *77*(September), 101726. https://doi.org/10.1016/j.evalprogplan.2019.101726

Paloş, R., Maricuţoiu, L. P., & Costea, I. (2019). Relations between academic performance, student engagement and student burnout: A cross-lagged analysis of a two-wave study. *Studies in Educational Evaluation*, *60*(September 2018), 199–204. https://doi.org/10.1016/j.stueduc.2019.01.005

Paluta, L., Kaiser, M. L., Huber-Krum, S., & Wheeler, J. (2019). Evaluating the impact of a healthy corner store initiative on food access domains. *Evaluation and Program Planning*, *73*(March 2018), 24–32. https://doi.org/10.1016/j.evalprogplan.2018.11.004

Parekh, J., Blum, R., Caldas, V., Whitfield, B., & Jennings, J. M. (2019). Program implementer perspectives replicating evidence based sexual reproductive health programs. *Evaluation and Program Planning*, *73*(July 2018), 80–87. https://doi.org/10.1016/j.evalprogplan.2018.12.004

Parrott, A., & Carman, J. G. (2019). Scaling Up Programs: Reflections on the Importance of Process Evaluation. *Canadian Journal of Program Evaluation*, *34*(1), 131–138. https://doi.org/10.3138/cjpe.43216

Patel, M. (2019). Anne Vo and Christina A. Christie (Eds.). (2015). Evaluation Use and Decision Making in Society: A Tribute to Marvin C. Alkin. Charlotte, NC: Information Age Publishing. *Canadian Journal of Program Evaluation*, *34*(2), 385–387. https://doi.org/10.3138/cjpe.67854

Patton, M. Q. (2019). Book Review: Measuring what matters in agriculture and food systems: A synthesis of the results and recommendations of TEEB for Agriculture and Food's Scientific and Economic Foundations report and TEEB for agriculture & food: Scientific and economic foun. *American Journal of Evaluation*, *40*(3), 459–464. https://doi.org/10.1177/1098214019834073

Patton, M. Q. (2019). Proposal for Evaluation of Women Affirming Motherhood. *New Directions for Evaluation*, *2019*(163), 27–47. https://doi.org/10.1002/ev.20375

Patton, M. Q. (2019). Transformation to Global Sustainability: Implications for Evaluation and Evaluators. *New Directions for Evaluation*, *2019*(162), 103–117. https://doi.org/10.1002/ev.20362

Patton, M. Q. (2019). On the Life and Death Importance of Thinking. *American Journal of Evaluation*, *40*(1), 137–146. https://doi.org/10.1177/1098214018756579

Pattyn, V., Molenveld, A., & Befani, B. (2019). Qualitative Comparative Analysis as an Evaluation Tool: Lessons From an Application in Development Cooperation. *American Journal of Evaluation*, *40*(1), 55–74. https://doi.org/10.1177/1098214017710502

Pawlowski, C. S., Schmidt, T., Nielsen, J. V., Troelsen, J., & Schipperijn, J. (2019). Will the children use it?—A RE-AIM evaluation of a local public open space intervention involving children from a deprived neighbourhood. *Evaluation and Program Planning*, *77*(August), 101706. https://doi.org/10.1016/j.evalprogplan.2019.101706

Pelucha, M., Kveton, V., & Potluka, O. (2019). Using mixed method approach in measuring effects of training in firms: Case study of the European Social Fund support. *Evaluation and Program Planning*, *73*(August 2018), 146–155. https://doi.org/10.1016/j.evalprogplan.2018.12.008

Perrin, B. (2019). How to Manage Pressure to Change Reports: Should Evaluators Be Above Criticism? *American Journal of Evaluation*, *40*(3), 354–375. https://doi.org/10.1177/1098214018792622

Persaud, N. (2019). An Exploratory Study on Public Sector Program Evaluation Practices and Culture in Barbados , Belize , Guyana , and Saint Vincent and the Grenadines : Where Are We ? Where Do We Need To Go ? *Journal of MultiDisciplinary Evaluation*, *15*(32), 17–27.

Peterman, K., & Gathings, M. J. (2019). Using a community-created multisite evaluation to promote evaluation use across a sector. *Evaluation and Program Planning*, *74*(February), 54–60. https://doi.org/10.1016/j.evalprogplan.2019.02.014

Peterson, C., & Skolits, G. (2019). Evaluating unintended program outcomes through Ripple Effects Mapping (REM): Application of REM using grounded theory. *Evaluation and Program Planning*, *76*(June), 101677. https://doi.org/10.1016/j.evalprogplan.2019.101677

Phelps, R. P. (2019). Test Frequency, Stakes, and Feedback in Student Achievement: A Meta-Analysis. *Evaluation Review*, *43*(3–4), 111–151. https://doi.org/10.1177/0193841X19865628

Phillips, G., Lindeman, P., Adames, C. N., Bettin, E., Bayston, C., Stonehouse, P., … Greene, G. J. (2019). Empowerment Evaluation: A Case Study of Citywide Implementation Within an HIV Prevention Context. *American Journal of Evaluation*, *40*(3), 318–334. https://doi.org/10.1177/1098214018796991

Picciotto, R. (2019). Donald T . Campbell's Evolutionary Perspective and its Implications for Evaluation. *Journal of MultiDisciplinary Evaluation*.

Picciotto, R. (2019). Is evaluation obsolete in a post-truth world? *Evaluation and Program Planning*, *73*(December 2018), 88–96. https://doi.org/10.1016/j.evalprogplan.2018.12.006

Picciotto, R. (2019). Is Adversary Evaluation Worth a Second Look? *American Journal of Evaluation*, *40*(1), 92–103. https://doi.org/10.1177/1098214018783068

Pierobon, C. (2019). EU's support to civil society in Kazakhstan: A pilot evaluation of the social capital generated. *Evaluation*, *25*(2), 207–223. https://doi.org/10.1177/1356389018796023

Podems, D. R. (2019). Book review. *Evaluation and Program Planning*, *77*, 101701. https://doi.org/10.1016/j.evalprogplan.2019.101701

Pop, C., & Khampirat, B. (2019). Self-assessment instrument to measure the competencies of Namibian graduates: Testing of validity and reliability. *Studies in Educational Evaluation*, *60*(January), 130–139. https://doi.org/10.1016/j.stueduc.2018.12.004

Porter, S., & Hawkins, P. (2019). Achieving Sustainability Through Sustainable Organizational Evaluation Systems. *New Directions for Evaluation*, *2019*(162), 87–101. https://doi.org/10.1002/ev.20360

Powell, S. (2019). Supplement to Theories of Change: Making Value Explicit. *Journal of Multidisciplinary Evaluation*, *15*(32), 53–54.

Powell, S. (2019). Theories of Change : Making Value Explicit. *Journal of MultiDisciplinary Evaluation*, *15*(32), 37–52.

Prilop, C. N., Weber, K. E., & Kleinknecht, M. (2019). How digital reflection and feedback environments contribute to pre-service teachers' beliefs during a teaching practicum. *Studies in Educational Evaluation*, *62*(June), 158–170. https://doi.org/10.1016/j.stueduc.2019.06.005

Pulido-Martos, M., Meléndez-Domínguez, M., & Lopez-Zafra, E. (2019). Cultural Adaptation and Psychometric Properties of the Shirom–Melamed Vigor Measure (SMVM) With Workers in Spain. *Evaluation and the Health Professions*, *42*(2), 219–232. https://doi.org/10.1177/0163278717734283

Quadrelli, C., Gould, L., & Rossingh, B. (2019). Editorial. *Evaluation Journal of Australasia*, *19*(2), 61–63. https://doi.org/10.1177/1035719X19858474

Quintelier, A., Vanhoof, J., & De Maeyer, S. (2019). A full array of emotions: An exploratory mixed methods study of teachers' emotions during a school inspection visit. *Studies in Educational Evaluation*, *63*(February), 83–93. https://doi.org/10.1016/j.stueduc.2019.07.006

Rabie, B., & Burger, A. (2019). Benefits of transport subsidisation: Comparing findings from a customer perception survey and Most Significant Change Technique interviews. *African Evaluation Journal*, *7*(1), 1–7. https://doi.org/10.4102/aej.v7i1.371

Ràfols, I. (2019). S&T indicators in the wild: Contextualization and participation for responsible metrics. *Research Evaluation*, *28*(1), 7–22. https://doi.org/10.1093/reseval/rvy030

Ramírez, R. (2019). Communication for Social Change: Seldom a Stand Alone, and Rarely Verified. *Journal of MultiDisciplinary Evaluation*, *15*(32), 1–16.

Ravn, R. (2019). Testing mechanisms in large-N realistic evaluations. *Evaluation*, *25*(2), 171–188. https://doi.org/10.1177/1356389019829164

Reddy, L. A., Hua, A., Dudek, C. M., Kettler, R. J., Lekwa, A., Arnold-Berkovits, I., & Crouse, K. (2019). Use of observational measures to predict student achievement. *Studies in Educational Evaluation*, *62*(May), 197–208. https://doi.org/10.1016/j.stueduc.2019.05.001

Renger, R., Atkinson, L., Renger, J., Renger, J., & Hart, G. (2019). The connection between logic models and systems thinking concepts. *Evaluation Journal of Australasia*, *19*(2), 79–87. https://doi.org/10.1177/1035719X19853660

Ricciuti, E., & Bufali, M. V. (2019). The health and social impact of Blood Donors Associations: A Social Return on Investment (SROI) analysis. *Evaluation and Program Planning*, *73*(June 2018), 204–213. https://doi.org/10.1016/j.evalprogplan.2019.01.003

Richardson, C. B., Morgan, N. R., Bleser, J. A., Aronson, K. R., & Perkins, D. F. (2019). A novel approach for evaluating programs designed to serve military veterans: Using an adapted common components analysis. *Evaluation and Program Planning*, *72*(September 2018), 145–151. https://doi.org/10.1016/j.evalprogplan.2018.10.012

Richardson, E. Z. L., Phillips, M., Colom, A., & Nichols, J. (2019). Honoring Lived Experience: Life Histories as a Realist Evaluation Method. *American Journal of Evaluation*, *40*(4), 509–532. https://doi.org/10.1177/1098214018796319

Richards-Schuster, K., & Plachta Elliott, S. (2019). A Practice Matrix for Involving Young People in Evaluation: Possibilities and Considerations. *American Journal of Evaluation*, *40*(4), 533–547. https://doi.org/10.1177/1098214019832113

Rickles, J., Zeiser, K. L., Yang, R., O'Day, J., & Garet, M. S. (2019). Promoting Deeper Learning in High School: Evidence of Opportunities and Outcomes. *Educational Evaluation and Policy Analysis*, *41*(2), 214–234. https://doi.org/10.3102/0162373719837949

Risien, J. (2019). Curators and sojourners in learning networks: Practices for transformation. *Evaluation and Program Planning*, *73*(September 2018), 71–79. https://doi.org/10.1016/j.evalprogplan.2018.12.001

Robinson-Garcia, N., Torres-Salinas, D., Herrera-Viedma, E., & Docampo, D. (2019). Mining university rankings: Publication output and citation impact as their basis. *Research Evaluation*, *28*(3), 232–240. https://doi.org/10.1093/reseval/rvz014

Rodriguez, R., Svensson, G., & Otero-Neira, C. (2019). Framing sustainable development through descriptive determinants in private hospitals – Orientation and organization. *Evaluation and Program Planning*, *75*(September 2018), 78–88. https://doi.org/10.1016/j.evalprogplan.2019.05.006

Rodríguez-Mantilla, J. M., Fernández-Díaz, M. J., & León Carrascosa, V. (2019). Validation of a questionnaire to evaluate the impact of ISO 9001 Standards in schools with a Confirmatory Factor Analysis. *Studies in Educational Evaluation*, *62*(May), 37–48. https://doi.org/10.1016/j.stueduc.2019.03.013

Roelofs, S., Edwards, N., Viehbeck, S., & Anderson, C. (2019). Formative, embedded evaluation to strengthen interdisciplinary team science: Results of a 4-year, mixed methods, multi-country case study. *Research Evaluation*, *28*(1), 37–50. https://doi.org/10.1093/reseval/rvy023

Rogers, A. (2019). Book review: The great nonprofit evaluation reboot: A new approach every staff member can understand. *Evaluation Journal of Australasia*, *19*(3), 151–153. https://doi.org/10.1177/1035719x19871085

Rogers, A. F., Kelly, L. M., & McCoy, A. (2019). Pathways to becoming an internal evaluator: Perspectives from the Australian non-government sector. *Evaluation and Program Planning*, *74*(January), 102–109. https://doi.org/10.1016/j.evalprogplan.2019.01.007

Rogers, A., Mccoy, A., & Kelly, L. M. (2019). Evaluation Literacy : Perspectives of Internal Evaluators in Non-Government Organizations. *Canadian Journal of Program Evaluation*, *1*, 1–20. https://doi.org/10.3138/cjpe.42190

Rolfe, S. (2019). Combining Theories of Change and Realist Evaluation in practice: Lessons from a research on evaluation study. *Evaluation*, *25*(3), 294–316. https://doi.org/10.1177/1356389019835229

Roorda, M., & Gullickson, A. M. (2019). Developing evaluation criteria using an ethical lens. *Evaluation Journal of Australasia*, *19*(4), 179–194. https://doi.org/10.1177/1035719X19891991

Rosval, L., Yamin, S., Jamshidi, P., & Aubry, T. (2019). Perceptions of the Use of an Outcome Monitoring Tool in a Clinical Psychology Training Centre: Lessons Learned for Performance Measurement. *Canadian Journal of Program Evaluation*, *34*(1), 84–101. https://doi.org/10.3138/cjpe.52823

Rotter, T., Plishka, C., Lawal, A., Harrison, L., Sari, N., Goodridge, D., … Kinsman, L. (2019). What Is Lean Management in Health Care? Development of an Operational Definition for a Cochrane Systematic Review. *Evaluation and the Health Professions*, *42*(3), 366–390. https://doi.org/10.1177/0163278718756992

Rowe, A. (2019). Rapid impact evaluation. *Evaluation*, *25*(4), 496–513. https://doi.org/10.1177/1356389019870213

Rowe, A. (2019). Sustainability-Ready Evaluation: A Call to Action. *New Directions for Evaluation*, *2019*(162), 29–48. https://doi.org/10.1002/ev.20365

Rozenszajn, R., Snapir, Z., & Machluf, Y. (2019). Professional learning and development of two groups of pre-service teachers with different scientific knowledge bases and different teaching training in the course of their studies. *Studies in Educational Evaluation*, *61*(March), 123–137. https://doi.org/10.1016/j.stueduc.2019.03.007

Ruel-Bergeron, J. C., Hurley, K. M., Kang, Y., Aburto, N., Farhikhtah, A., Dinucci, A., … Christian, P. (2019). Monitoring and evaluation design of Malawi's Right Foods at the Right Time nutrition program. *Evaluation and Program Planning*, *73*(February 2018), 1–9. https://doi.org/10.1016/j.evalprogplan.2018.11.001

Russon, C. (2019). Methods for Collecting Data Based Upon an Eastern Paradigm of Evaluation. *Journal of MultiDisciplinary Evaluation*, *15*(32), 59–63.

Ruvalcaba, E., Callaghan-Koru, J., Rand, C. S., & Eakin, M. N. (2019). Integrating asthma management and care in Maryland Head Start Programs: A qualitative assessment of opportunities and needs. *Evaluation and Program Planning*, *77*(July), 101684. https://doi.org/10.1016/j.evalprogplan.2019.101684

Saavedra, J. E., Näslund-Hadley, E., & Alfonso, M. (2019). Remedial Inquiry-Based Science Education: Experimental Evidence From Peru. *Educational Evaluation and Policy Analysis*, *41*(4), 483–509. https://doi.org/10.3102/0162373719867081

Sætnan, A. R., Tøndel, G., & Rasmussen, B. (2019). Does counting change what is counted? Potential for paradigm change through performance metrics. *Research Evaluation*, *28*(1), 73–83. https://doi.org/10.1093/reseval/rvy032

Sansone, D. (2019). Teacher Characteristics, Student Beliefs, and the Gender Gap in STEM Fields. *Educational Evaluation and Policy Analysis*, *41*(2), 127–144. https://doi.org/10.3102/0162373718819830

Savy, P., Hodgkin, S., Long, K., Melis, R., Lewis, V., Furler, J., … Blackberry, I. (2019). Comprehensive geriatric assessment programmes: Possibilities, realities and outcomes. *Evaluation Journal of Australasia*, *19*(3), 115–133. https://doi.org/10.1177/1035719X19876153

Schuller, K. A., Cronin, C. E., Nicks, S. E., Jing, X., Kingori, C., & Morrone, M. (2019). Development and application of a rubric to compare strategies for improving access to health care in rural communities in the United States. *Evaluation and Program Planning*, *74*(September 2018), 61–68. https://doi.org/10.1016/j.evalprogplan.2019.02.013

Schwandt, T. A. (2019). Post-normal evaluation? *Evaluation*, *25*(3), 317–329. https://doi.org/10.1177/1356389019855501

Schwarzman, J., Bauman, A., Gabbe, B. J., Rissel, C., Shilton, T., & Smith, B. J. (2019). Understanding the factors that influence health promotion evaluation: The development and validation of the evaluation practice analysis survey. *Evaluation and Program Planning*, *74*(March), 76–83. https://doi.org/10.1016/j.evalprogplan.2019.03.002

Scott, P. W. (2019). Causal Inference Methods for selection on observed and unobserved factors: Propensity Score Matching, Heckit Models, and Instrumental Variable Estimation. *Practical Assessment, Research and Evaluation*, *24*(3), 1–15.

Scott-Clayton, J., & Wen, Q. (2019). Estimating Returns to College Attainment: Comparing Survey and State Administrative Data–Based Estimates. *Evaluation Review*, *43*(5), 266–306. https://doi.org/10.1177/0193841X18803247

Scriven, M. (2019). The Checklist Imperative. *New Directions for Evaluation*, *2019*(163), 49–60. https://doi.org/10.1002/ev.20374

Sellick, S. (2019). Collaborative, Participatory, and Empowerment Evaluation: Stakeholder Involvement Approaches, by David M. Fetterman, Liliana Rodríguez-Campos, Ann P. Zukoski, et al. *Canadian Journal of Program Evaluation*, *34*(1), 168–170. https://doi.org/10.3138/cjpe.56975

Seo, Y. J., & Um, K. H. (2019). The asymmetric effect of fairness and quality dimensions on satisfaction and dissatisfaction: An application of the Kano model to the interdisciplinary college program evaluation. *Studies in Educational Evaluation*, *61*(April), 183–195. https://doi.org/10.1016/j.stueduc.2019.04.007

Shannon Guss, C. E. B. (2019). Using Program Theory to Evaluate a Graduate Student Development Program. *Journal of MultiDisciplinary Evaluation*, *15*(33), 66–80. Retrieved from http://journals.sfu.ca/jmde/index.php/jmde_1/article/view/559/497

Shapiro, A., & Weiland, C. (2019). What Is in a Definition? The How and When of Special Education Subgroup Analysis in Preschool Evaluations. *Educational Evaluation and Policy Analysis*, *41*(2), 145–163. https://doi.org/10.3102/0162373718820307

Sheppard, C., Baker, A., Lolic, P., Soni, S., & Courtney, B. (2019). 20 Years Later: Reflections on the CES Student Evaluation Case Competition. *Canadian Journal of Program Evaluation*, *34*(1), 118–130. https://doi.org/10.3138/cjpe.42239

Shin, H., Lee, S. J., Lee, Y. nah, & Shon, S. (2019). Community health needs assessment for a child health promotion program in Kyrgyzstan. *Evaluation and Program Planning*, *74*(February), 1–9. https://doi.org/10.1016/j.evalprogplan.2019.02.005

Shmueli, D. F., Ben Gal, M., Segal, E., Reichman, A., & Feitelson, E. (2019). How can regulatory systems be assessed? The case of earthquake preparedness in Israel. *Evaluation*, *25*(1), 80–98. https://doi.org/10.1177/1356389018803235

Siddiqui, N., Gorard, S., & See, B. H. (2019). Can learning beyond the classroom impact on social responsibility and academic attainment? An evaluation of the Children's University youth social action programme. *Studies in Educational Evaluation*, *61*(March), 74–82. https://doi.org/10.1016/j.stueduc.2019.03.004

Siebert, P., & Myles, P. (2019). Eliciting and reconstructing programme theory: An exercise in translating theory into practice. *Evaluation*, *25*(4), 469–476. https://doi.org/10.1177/1356389019870211

Singletary, C. R., Weaver, G., Carson, R. L., Beets, M. W., Pate, R. R., Saunders, R. P., … Moore, J. B. (2019). Evaluation of a comprehensive school physical activity program: Be a Champion! *Evaluation and Program Planning*, *75*(May), 54–60. https://doi.org/10.1016/j.evalprogplan.2019.04.003

Solans-Domènech, M., MV Pons, J., Adam, P., Grau, J., & Aymerich, M. (2019). Development and validation of a questionnaire to measure research impact. *Research Evaluation*, *28*(3), 253–262. https://doi.org/10.1093/reseval/rvz007

Sridharan, S., & Nakaima, A. (2019). Till time (and poor planning) do us part: Programs as dynamic systems—Incorporating planning of sustainability into theories of change.

Canadian Journal of Program Evaluation, *33*(3), 375–394. https://doi.org/10.3138/cjpe.53055

Stame, N. (2019). Doubt, surprise, and the ethical evaluator: Lessons from the work of Judith Tendler. *Evaluation*, *25*(4), 449–468. https://doi.org/10.1177/1356389019879458

Stern, E. (2019). Editorial. *Evaluation*, *25*(2), 127–130. https://doi.org/10.1177/1356389019842679

Stern, E. (2019). Editorial. *Evaluation*, *25*(1), 3–5. https://doi.org/10.1177/1356389018823110

Stern, E. (2019). Editorial. *Evaluation*, *25*(4), 409–410. https://doi.org/10.1177/1356389019881046

Stern, E. (2019). Editorial. *Evaluation*, *25*(3), 257–260. https://doi.org/10.1177/1356389019862682

Stewart, G., Webster, C. A., Weaver, R. G., Stodden, D. F., Brian, A., Egan, C. A., … Patey, M. (2019). Evaluation of a classroom movement integration training delivered in a low socioeconomic school district. *Evaluation and Program Planning*, *73*(August 2018), 187–194. https://doi.org/10.1016/j.evalprogplan.2018.12.010

Stickl Haugen, J., & Chouinard, J. A. (2019). Transparent, Translucent, Opaque: Exploring the Dimensions of Power in Culturally Responsive Evaluation Contexts. *American Journal of Evaluation*, *40*(3), 376–394. https://doi.org/10.1177/1098214018796342

Stockdale, J., Hughes, C., Stronge, S., & Birch, M. (2019). Motivating midwifery students to digitalise their enquiry-based learning experiences: An evaluative case study. *Studies in Educational Evaluation*, *60*(June 2018), 59–65. https://doi.org/10.1016/j.stueduc.2018.11.006

Stone-Jovicich, S., Percy, H., McMillan, L., Turner, J. A., Chen, L., & White, T. (2019). Evaluating monitoring, evaluation and learning initiatives in the New Zealand and Australian agricultural research and innovation systems: The MEL 2 framework. *Evaluation Journal of Australasia*, *19*(1), 8–21. https://doi.org/10.1177/1035719X18823567

Striebing, C., Kalpazidou Schmidt, E., & Palmén, R. (2019). Pragmatic ex-ante evaluation using an innovative conceptual framework: The case of a high-tech entrepreneurship program for women. *Evaluation and Program Planning*, *77*(May), 101714. https://doi.org/10.1016/j.evalprogplan.2019.101714

Stura, I., Gentile, T., Migliaretti, G., & Vesce, E. (2019). Accreditation in higher education: Does disciplinary matter? *Studies in Educational Evaluation*, *63*(November 2018), 41–47. https://doi.org/10.1016/j.stueduc.2019.07.004

Sun, M., Liu, J., Zhu, J., & LeClair, Z. (2019). Using a Text-as-Data Approach to Understand Reform Processes: A Deep Exploration of School Improvement Strategies. *Educational Evaluation and Policy Analysis*, *41*(4), 510–536. https://doi.org/10.3102/0162373719869318

Suranga, S. S., Senadhira, K., & Rajakaruna, R. M. D. K. (2019). The branch performance tool: A methodological approach for continuous monitoring of programme efficiencies of service delivery interventions. *African Evaluation Journal*, *7*(1), 1–8. https://doi.org/10.4102/aej.v7i1.262

Sussman, J., & Wilson, M. R. (2019). The Use and Validity of Standardized Achievement Tests for Evaluating New Curricular Interventions in Mathematics and Science. *American Journal of Evaluation*, *40*(2), 190–213. https://doi.org/10.1177/1098214018767313

Sweeney, A. M., Wilson, D. K., & Brown, A. (2019). A qualitative study to examine how differences in motivation can inform the development of targeted physical activity interventions for African American women. *Evaluation and Program Planning*, *77*(March), 101718. https://doi.org/10.1016/j.evalprogplan.2019.101718

Szijarto, B., & Bradley Cousins, J. (2019). Mapping the practice of developmental evaluation: Insights from a concept mapping study. *Evaluation and Program Planning*, *76*(February), 101666. https://doi.org/10.1016/j.evalprogplan.2019.05.001

Szijarto, B., & Cousins, J. B. (2019). Making Space for Adaptive Learning. *American Journal of Evaluation*, *40*(2), 160–176. https://doi.org/10.1177/1098214018781506

Tarsilla, M. (2019). Scott G. Chaplowe and J. Bradley Cousins. (2016). Monitoring and Evaluation Training: A Systematic Approach. Thousand Oaks, CA: SAGE. *Canadian Journal of Program Evaluation*, *34*(2), 388–390. https://doi.org/10.3138/cjpe.67892

Tavana, M., Khosrojerdi, G., Mina, H., & Rahman, A. (2019). A hybrid mathematical programming model for optimal project portfolio selection using fuzzy inference system and analytic hierarchy process. *Evaluation and Program Planning*, *77*(August), 101703. https://doi.org/10.1016/j.evalprogplan.2019.101703

Taylor, J. M. (2019). Overview and illustration of Bayesian confirmatory factor analysis with ordinal indicators. *Practical Assessment, Research and Evaluation*, *24*(4), 1–16.

Taylor, T., & Scott, A. (2019). Do Physicians Prefer to Complete Online or Mail Surveys? Findings From a National Longitudinal Survey. *Evaluation and the Health Professions*, *42*(1), 41–70. https://doi.org/10.1177/0163278718807744

ter Beek, M., Opdenakker, M. C., Deunk, M. I., & Strijbos, J. W. (2019). Teaching reading strategies in history lessons: A micro-level analysis of professional development training and its practical challenges. *Studies in Educational Evaluation*, *63*(May), 26–40. https://doi.org/10.1016/j.stueduc.2019.07.003

Testa, G., Szkuta, K., & Cunningham, P. N. (2019). Improving access to finance for young innovative enterprises with growth potential: Evidence of impact of R&D grant schemes on firms' outputs. *Research Evaluation*, *28*(4), 355–369. https://doi.org/10.1093/reseval/rvz016

Theresa Hunter, M., & McDavid, J. C. (2019). Comparison of Canadian and American graduate evaluation education programs. *Canadian Journal of Program Evaluation*, *34*(2), 207–234. https://doi.org/10.3138/cjpe.56989

Tipton, E., & Matlen, B. J. (2019). Improved Generalizability Through Improved Recruitment: Lessons Learned From a Large-Scale Randomized Trial. *American Journal of Evaluation*, *40*(3), 414–430. https://doi.org/10.1177/1098214018810519

Tong, F., Tang, S., Irby, B. J., Lara-Alecio, R., Guerrero, C., & Lopez, T. (2019). A process for establishing and maintaining inter-rater reliability for two observation instruments as a fidelity of implementation measure: A large-scale randomized controlled trial perspective. *Studies in Educational Evaluation*, *62*(April), 18–29. https://doi.org/10.1016/j.stueduc.2019.04.008

Trinidad, J. E. (2019). Understanding when parental aspirations negatively affect student outcomes: The case of aspiration-expectation inconsistency. *Studies in Educational Evaluation*, *60*(June 2018), 179–188. https://doi.org/10.1016/j.stueduc.2019.01.004

Tsang, A., Fung, D., & Yau, A. H. Y. (2019). Evaluating supplementary and mainstream ESL/EFL education: Learners' views from secondary- and tertiary-level perspectives. *Studies in Educational Evaluation*, *62*(May), 61–71. https://doi.org/10.1016/j.stueduc.2019.04.004

Tsey, K., Onnis, L. ann, Whiteside, M., McCalman, J., Williams, M., Heyeres, M., … Kinchin, I. (2019). Assessing research impact: Australian Research Council criteria and the case of Family Wellbeing research. *Evaluation and Program Planning*, *73*(January), 176–186. https://doi.org/10.1016/j.evalprogplan.2019.01.004

Turner, B. (2019). Book reviews: The Sage qualitative research kit: Analyzing qualitative data. *Evaluation Journal of Australasia*, *19*(1), 56–57. https://doi.org/10.1177/1035719x19827672

Uitto, J. I. (2019). Sustainable Development Evaluation: Understanding the Nexus of Natural and Human Systems. *New Directions for Evaluation*, *2019*(162), 49–67. https://doi.org/10.1002/ev.20364

Valero-Mora, P., Rodrigo, M. F., Sanchez, M., & Sanmartin, J. (2019). A plot for the visualization of missing value patterns in multivariate data. *Practical Assessment, Research and Evaluation*, *24*(6). https://doi.org/10.7275/94ra-1y55

Veletanlić, E., & Sá, C. (2019). Government programs for university-industry partnerships: Logics, design, and implications for academic science. *Research Evaluation*, *28*(2), 109–122. https://doi.org/10.1093/reseval/rvy034

Vera-Villarroel, P., Urzúa, A., Jaime, D., Contreras, D., Zych, I., Celis-Atenas, K., … Lillo, S. (2019). Positive and Negative Affect Schedule (PANAS): Psychometric Properties and Discriminative Capacity in Several Chilean Samples. *Evaluation and the Health Professions*, *42*(4), 473–497. https://doi.org/10.1177/0163278717745344

Vichta, R. (2019). Book reviews: Research proposal: Little quick fix. *Evaluation Journal of Australasia*, *19*(1), 54–56. https://doi.org/10.1177/1035719x19827671

Wadsworth, Y. (2019). Evaluation perspectives. *Evaluation Journal of Australasia*, *19*(2), 101–106. https://doi.org/10.1177/1035719X19857507

Walker, J. T., Salter, A., Fontinha, R., & Salandra, R. (2019). The impact of journal re-grading on perception of ranking systems: Exploring the case of the academic journal guide and business and management scholars in the UK. *Research Evaluation*, *28*(3), 218–231. https://doi.org/10.1093/reseval/rvz010

Wang, C. Y., Chen, J. De, Wang, C. H., Wang, J. Y., Tai, C. J., Hsieh, T. Y., & Chen, D. Y. (2019). Effects of Job Burnout and Emotional Labor on Objective Structured Clinical

Examination Performance Among Interns and Residents in Taiwan. *Evaluation and the Health Professions*, *42*(2), 233–257. https://doi.org/10.1177/0163278717729732

Wang, S. R. (2019). Positive Practice Toward Voluntary Unpaid Blood Donation Among College Students in Shandong, China, 2013–2017. *Evaluation and the Health Professions*, *42*(1), 114–117. https://doi.org/10.1177/0163278718788927

Wang-Schweig, M., Miller, B. A., Buller, D. B., Byrnes, H. F., Bourdeau, B., & Rogers, V. (2019). Using Panel Vendors for Recruitment Into a Web-Based Family Prevention Program: Methodological Considerations. *Evaluation and the Health Professions*, *42*(1), 24–40. https://doi.org/10.1177/0163278717742189

Ward, R. C., Muckle, T. J., Kremer, M. J., & Krogh, M. A. (2019). Computer-Based Case Simulations for Assessment in Health Care: A Literature Review of Validity Evidence. *Evaluation and the Health Professions*, *42*(1), 82–102. https://doi.org/10.1177/0163278717718609

Warinda, E. (2019). Evaluating operationalisation of integrated monitoring and evaluation system in Kisumu County: Implications for policy makers. *African Evaluation Journal*, *7*(1), 1–12. https://doi.org/10.4102/aej.v7i1.385

Wehipeihana, N. (2019). Increasing cultural competence in support of indigenous-led evaluation: A necessary step toward indigenous-led evaluation. *Canadian Journal of Program Evaluation*, *34*(2), 368–384. https://doi.org/10.3138/cjpe.68444

Were, L., Crocket, A., McKegg, K., & Goodwin, D. (2019). Review of Principles-Focused Evaluation: The GUIDE . *American Journal of Evaluation*, *40*(2), 306–310. https://doi.org/10.1177/1098214018785468

Whynot, J., Montague, S., & Lemire, S. (2019). The current landscape of program theorizing. *Canadian Journal of Program Evaluation*, *33*(3), 287–291. https://doi.org/10.3138/cjpe.56900

Wiley, C. (2019). Standardised module evaluation surveys in UK higher education: Establishing students' perspectives. *Studies in Educational Evaluation*, *61*(February), 55–65. https://doi.org/10.1016/j.stueduc.2019.02.004

Wind, S. A., Jones, E., Bergin, C., & Jensen, K. (2019). Exploring patterns of principal judgments in teacher evaluation related to reported gender and years of experience. *Studies*

in Educational Evaluation, *61*(March), 150–158. https://doi.org/10.1016/j.stueduc.2019.03.011

Wolfe, S. M., & Price, A. W. (2019). Partnering for Social Change: Collaborating With Clients and Communities. *New Directions for Evaluation*, *2019*(164), 89–100. https://doi.org/10.1002/ev.20384

Wu, H., Shen, J., Jones, J., Gao, X., Zheng, Y., & Krenn, H. Y. (2019). Using logic model and visualization to conduct portfolio evaluation. *Evaluation and Program Planning*, *74*(September 2018), 69–75. https://doi.org/10.1016/j.evalprogplan.2019.02.011

Wusinich, C., Bond, L., Nathanson, A., & Padgett, D. K. (2019). "If you're gonna help me, help me": Barriers to housing among unsheltered homeless adults. *Evaluation and Program Planning*, *76*(June), 101673. https://doi.org/10.1016/j.evalprogplan.2019.101673

Yao, D., Jiang, L., Huang, Y., Chen, L., Wang, Y., & Xi, X. (2019). Assessment of Factors Influencing Communication in Clinical Pharmacy. *Evaluation and the Health Professions*, *42*(4), 422–449. https://doi.org/10.1177/0163278717751028

Yasmin, M., Naseem, F., & Masso, I. C. (2019). Teacher-directed learning to self-directed learning transition barriers in Pakistan. *Studies in Educational Evaluation*, *61*(February), 34–40. https://doi.org/10.1016/j.stueduc.2019.02.003

Yoo, J. (2019). The impact of conflict among political actors on implementing South Korea's new teacher evaluation policy: A case study with implications for education policymaking. *Studies in Educational Evaluation*, *61*(March), 94–104. https://doi.org/10.1016/j.stueduc.2019.03.006

Youker, B. W. (2019). What, how, and why? A comparative analysis of 12 goal-free evaluations. *Journal of MultiDisciplinary Evaluation*, *15*(33), 16–29. Retrieved from http://www.jmde.com

Young, T. L., Janke, M. C., Sharpe, C., & Carthron, D. (2019). Evaluating the feasibility of a community intergenerational physical activity intervention for kinship families: Professional stakeholders' perspectives. *Evaluation and Program Planning*, *72*(September 2018), 136–144. https://doi.org/10.1016/j.evalprogplan.2018.09.002

Yousefi Nooraie, R., Khan, S., Gutberg, J., & Baker, G. R. (2019). A Network Analysis Perspective to Implementation: The Example of Health Links to Promote Coordinated Care.

Evaluation and the Health Professions, *42*(4), 395–421. https://doi.org/10.1177/0163278718772887

Yu, S., Zhou, N., Zheng, Y., Zhang, L., Cao, H., & Li, X. (2019). Evaluating student motivation and engagement in the Chinese EFL writing context. *Studies in Educational Evaluation*, *62*(19), 129–141. https://doi.org/10.1016/j.stueduc.2019.06.002

Yue, J. J., & Chen, G. (2019). An exploration of the structure of pharmaceutical mentors' competence at universities in Western China. *Studies in Educational Evaluation*, *63*(30), 1–8. https://doi.org/10.1016/j.stueduc.2019.06.004

Zacarías, G., Shamah-Levy, T., Elton-Puente, E., Garbus, P., & García, O. P. (2019). Development of an intervention program to prevent childhood obesity targeted to Mexican mothers of school-aged children using intervention mapping and social cognitive theory. *Evaluation and Program Planning*, *74*(February), 27–37. https://doi.org/10.1016/j.evalprogplan.2019.02.008

Zack, M. K., Karre, J. K., Olson, J., & Perkins, D. F. (2019). Similarities and differences in program registers: A case study. *Evaluation and Program Planning*, *76*(June), 101676. https://doi.org/10.1016/j.evalprogplan.2019.101676

Zala-Mezö, E., Raeder, S., & Strauss, N. C. (2019). More knowledge helps implement change: Evaluation of a training program during educational reform. *Studies in Educational Evaluation*, *61*(March), 105–111. https://doi.org/10.1016/j.stueduc.2019.03.008

Zhang, Y. P., Liu, W. H., Yan, Y. T., Porr, C., Zhang, Y., & Wei, H. H. (2019). Psychometric Testing of the Evidence-Based Practice Nursing Leadership Scale and the Work Environment Scale After Cross-Cultural Adaptation in Mainland China. *Evaluation and the Health Professions*, *42*(3), 328–343. https://doi.org/10.1177/0163278718801439

Zvoch, K. (2019). Investigation of the long term effect of a summer literacy program on student reading performance. *Studies in Educational Evaluation*, *62*(January 2018), 111–117. https://doi.org/10.1016/j.stueduc.2019.05.005

Appendix B

Evaluation Word List

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
evaluation	24,440	12,257	6,080.45	129.91	480	3,257	99.38	24.83	6.27
program	12,399	18,967	3,084.76	201.02	384	3,619	79.50	27.59	3.40
student	11,907	42,198	2,962.35	447.24	257	2,745	53.21	20.93	2.74
teacher	8,401	20,915	2,090.09	221.67	179	1,309	37.06	9.98	2.53
school	9,172	31,696	2,281.91	335.93	327	2,920	67.70	22.26	2.46
research	10,523	49,977	2,618.03	529.69	453	7,358	93.79	56.10	2.37
intervention	4,947	9,566	1,230.77	101.39	287	2,101	59.42	16.02	2.03
outcome	5,378	18,155	1,338.00	192.42	396	4,016	81.99	30.62	1.96
evaluator	3,884	470	966.30	4.98	244	87	50.52	0.66	1.96
project	5,306	22,335	1,320.08	236.72	337	4,041	69.77	30.81	1.88
assessment	4,466	12,880	1,111.10	136.51	353	3,307	73.08	25.22	1.86
health	4,381	17,446	1,089.95	184.90	300	2,697	62.11	20.56	1.76
impact	4,648	21,317	1,156.38	225.93	394	5,213	81.57	39.75	1.76
et	8,565	77,138	2,130.89	817.55	420	6,285	86.96	47.92	1.72
participant	5,170	31,088	1,286.25	329.49	344	2,908	71.22	22.17	1.72
implementation	3,187	7,849	792.90	83.19	315	2,184	65.22	16.65	1.66
community	4,646	28,802	1,155.88	305.26	351	3,786	72.67	28.87	1.65
education	4,031	22,977	1,002.88	243.52	369	2,710	76.40	20.66	1.61
programme	2,705	6,522	672.98	69.12	122	1,453	25.26	11.08	1.57
learn	3,968	25,435	987.20	269.57	377	3,838	78.05	29.26	1.57
stakeholder	2,406	2,194	598.59	23.25	273	474	56.52	3.61	1.56
team	2,917	10,067	725.72	106.70	276	1,424	57.14	10.86	1.56
practice	4,495	36,568	1,118.31	387.57	420	5,847	86.96	44.58	1.53
approach	5,178	47,740	1,288.24	505.98	449	8,537	92.96	65.09	1.52
item	3,368	19,933	837.93	211.26	193	2,415	39.96	18.41	1.52
study	11,041	140,050	2,746.90	1,484.33	448	11,421	92.75	87.08	1.51
how	5,767	59,079	1,434.78	626.15	466	8,824	96.48	67.28	1.50
survey	2,882	14,293	717.02	151.49	258	3,010	53.42	22.95	1.49
focus	4,145	34,286	1,031.24	363.38	457	8,039	94.62	61.30	1.49
university	2,761	12,532	686.91	132.82	398	3,901	82.40	29.74	1.49
score	3,403	22,988	846.64	243.64	222	2,755	45.96	21.01	1.49
researcher	2,555	10,551	635.66	111.83	330	3,035	68.32	23.14	1.47
development	4,535	42,973	1,128.27	455.45	417	7,352	86.34	56.06	1.46
support	4,636	45,979	1,153.39	487.31	438	8,860	90.68	67.56	1.45
need	5,096	53,660	1,267.84	568.72	460	9,860	95.24	75.18	1.45
professional	2,414	10,437	600.58	110.62	287	2,259	59.42	17.22	1.44
educational	2,222	7,297	552.81	77.34	272	1,533	56.31	11.69	1.44
datum	6,642	80,212	1,652.47	850.13	442	8,953	91.51	68.27	1.43

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
learning	2,320	9,613	577.19	101.88	261	1,471	54.04	11.22	1.43
interview	2,416	11,565	601.08	122.57	267	1,941	55.28	14.80	1.43
training	2,657	15,590	661.04	165.23	287	2,242	59.42	17.09	1.43
process	6,151	73,162	1,530.31	775.41	446	9,815	92.34	74.84	1.43
funding	1,837	2,477	457.03	26.25	238	818	49.28	6.24	1.42
policy	3,737	34,396	929.73	364.55	336	3,200	69.57	24.40	1.41
page	1,920	4,863	477.68	51.54	298	1,461	61.70	11.14	1.41
design	3,521	32,243	875.99	341.73	434	6,693	89.86	51.03	1.40
skill	2,334	12,315	580.68	130.52	312	2,340	64.60	17.84	1.40
theory	3,731	36,158	928.24	383.22	345	5,397	71.43	41.15	1.39
include	5,968	74,103	1,484.78	785.39	474	11,596	98.14	88.42	1.39
staff	1,938	6,380	482.16	67.62	276	1,465	57.14	11.17	1.39
review	2,480	15,671	617.00	166.09	394	4,933	81.57	37.61	1.39
knowledge	3,663	35,801	911.32	379.44	402	5,789	83.23	44.14	1.39
improve	2,526	16,492	628.45	174.79	396	5,521	81.99	42.10	1.39
science	2,218	11,333	551.82	120.11	265	2,354	54.87	17.95	1.39
finding	2,919	23,557	726.22	249.67	399	5,664	82.61	43.19	1.38
conduct	2,435	15,654	605.81	165.91	416	5,364	86.13	40.90	1.38
develop	3,755	38,343	934.21	406.38	441	8,849	91.30	67.47	1.38
feedback	1,829	6,481	455.04	68.69	228	1,262	47.20	9.62	1.36
framework	2,208	13,805	549.33	146.31	348	4,079	72.05	31.10	1.35
year	4,481	53,577	1,114.83	567.84	417	8,131	86.34	62.00	1.35
context	3,455	35,672	859.57	378.07	414	6,590	85.71	50.25	1.35
planning	1,828	7,516	454.79	79.66	260	1,297	53.83	9.89	1.35
academic	1,856	8,047	461.76	85.29	253	1,874	52.38	14.29	1.35
implement	1,899	9,586	472.45	101.60	338	3,222	69.98	24.57	1.34
provide	5,474	72,536	1,361.88	768.78	473	11,311	97.93	86.24	1.34
organization	2,447	19,350	608.79	205.08	292	3,097	60.46	23.61	1.34
indicator	1,752	7,218	435.88	76.50	230	2,314	47.62	17.64	1.33
identify	3,385	36,018	842.16	381.74	438	8,591	90.68	65.51	1.33
work	5,956	81,889	1,481.80	867.91	459	9,751	95.03	74.35	1.33
table	3,519	39,142	875.49	414.85	359	7,145	74.33	54.48	1.33
experience	3,446	38,734	857.33	410.53	419	6,018	86.75	45.89	1.32
service	2,827	27,650	703.33	293.05	281	3,626	58.18	27.65	1.32
classroom	1,584	5,495	394.08	58.24	116	679	24.02	5.18	1.32
achievement	1,540	4,902	383.14	51.95	189	1,486	39.13	11.33	1.32
journal	1,426	3,145	354.78	33.33	356	1,114	73.71	8.49	1.31
grade	1,561	5,683	388.36	60.23	127	1,436	26.29	10.95	1.31

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
quality	2,482	22,338	617.50	236.75	361	4,970	74.74	37.90	1.31
college	1,497	4,812	372.44	51.00	150	1,276	31.06	9.73	1.31
sustainability	1,293	1,290	321.69	13.67	113	289	23.40	2.20	1.30
change	5,671	80,168	1,410.89	849.67	429	10,401	88.82	79.31	1.30
strategy	2,788	28,621	693.63	303.34	347	5,273	71.84	40.21	1.30
author	1,957	13,657	486.88	144.74	432	4,821	89.44	36.76	1.30
assess	2,057	15,537	511.76	164.67	357	5,139	73.91	39.18	1.30
course	2,340	20,870	582.17	221.19	251	6,004	51.97	45.78	1.30
evaluate	1,986	14,545	494.10	154.16	374	5,151	77.43	39.28	1.30
participate	1,667	8,969	414.73	95.06	303	3,623	62.73	27.62	1.29
help	2,272	20,131	565.25	213.36	415	6,324	85.92	48.22	1.29
challenge	1,983	15,076	493.35	159.78	377	4,734	78.05	36.10	1.29
across	2,470	24,004	614.51	254.41	387	6,011	80.12	45.83	1.29
teaching	1,578	7,816	392.59	82.84	160	1,266	33.13	9.65	1.29
report	4,165	55,225	1,036.21	585.31	413	9,068	85.51	69.14	1.28
self	1,411	5,510	351.04	58.40	201	1,309	41.61	9.98	1.28
tool	1,654	10,140	411.50	107.47	321	3,582	66.46	27.31	1.28
initiative	1,306	4,012	324.92	42.52	236	1,398	48.86	10.66	1.27
address	1,984	16,769	493.60	177.73	421	6,036	87.16	46.02	1.27
engagement	1,265	3,997	314.72	42.36	248	1,245	51.35	9.49	1.26
a1	1,195	2,770	297.31	29.36	201	570	41.61	4.35	1.26
benefit	1,989	17,832	494.84	188.99	347	4,580	71.84	34.92	1.26
resource	2,090	19,935	519.97	211.28	364	3,904	75.36	29.77	1.26
analysis	5,200	78,628	1,293.71	833.35	439	10,461	90.89	79.76	1.25
goal	1,852	16,624	460.76	176.19	358	4,237	74.12	32.31	1.24
improvement	1,372	7,716	341.34	81.78	276	2,803	57.14	21.37	1.24
their	12,105	211,018	3,011.61	2,236.49	477	12,276	98.76	93.60	1.24
qualitative	1,153	3,739	286.86	39.63	245	1,581	50.72	12.05	1.24
group	5,901	93,837	1,468.11	994.54	432	9,051	89.44	69.01	1.24
plan	1,697	14,174	422.20	150.22	299	3,115	61.90	23.75	1.24
capacity	1,670	13,759	415.48	145.83	256	3,594	53.00	27.40	1.24
validity	1,247	5,922	310.24	62.76	214	2,011	44.31	15.33	1.23
collaboration	1,047	2,438	260.48	25.84	202	877	41.82	6.69	1.23
activity	3,623	51,653	901.37	547.45	372	7,103	77.02	54.16	1.23
data	1,564	12,313	389.11	130.50	344	4,078	71.22	31.09	1.23
district	1,169	4,757	290.84	50.42	97	926	20.08	7.06	1.23
social	3,871	56,852	963.07	602.55	406	4,756	84.06	36.26	1.23
partner	1,268	7,281	315.47	77.17	176	1,654	36.44	12.61	1.22

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
gender	1,414	10,258	351.79	108.72	178	1,752	36.85	13.36	1.22
partnership	1,024	2,793	254.76	29.60	161	647	33.33	4.93	1.22
article	2,074	23,052	515.99	244.32	338	4,587	69.98	34.98	1.22
opportunity	1,597	14,100	397.32	149.44	350	4,194	72.46	31.98	1.22
receive	1,870	19,439	465.24	206.03	366	5,969	75.78	45.51	1.22
key	1,802	18,119	448.32	192.04	375	5,979	77.64	45.59	1.22
engage	1,387	10,402	345.07	110.25	309	3,476	63.98	26.50	1.21
management	1,850	19,523	460.26	206.92	272	2,859	56.31	21.80	1.21
evidence	2,722	36,488	677.21	386.72	376	7,520	77.85	57.34	1.21
criterion	1,510	12,969	375.67	137.45	268	3,884	55.49	29.61	1.21
questionnaire	1,118	5,374	278.15	56.96	145	1,219	30.02	9.29	1.21
question	3,113	44,286	774.49	469.37	425	7,677	87.99	58.54	1.21
use	14,048	258,839	3,495.01	2,743.33	479	12,977	99.17	98.95	1.20
attitude	1,387	11,290	345.07	119.66	184	2,499	38.10	19.05	1.20
member	2,299	29,207	571.97	309.55	301	5,141	62.32	39.20	1.20
about	5,482	91,810	1,363.87	973.06	464	10,467	96.07	79.81	1.20
youth	1,033	4,775	257.00	50.61	75	846	15.53	6.45	1.20
level	5,061	84,353	1,259.13	894.02	438	10,198	90.68	77.76	1.19
child	3,030	44,401	753.84	470.59	194	3,033	40.17	23.13	1.19
public	2,435	32,607	605.81	345.59	335	4,153	69.36	31.67	1.19
peer	1,017	4,735	253.02	50.18	166	981	34.37	7.48	1.19
doi	778	156	193.56	1.65	190	37	39.34	0.28	1.19
field	1,990	24,023	495.09	254.61	327	5,228	67.70	39.86	1.19
literature	1,725	18,735	429.16	198.56	366	5,773	75.78	44.02	1.19
online	1,019	4,904	253.52	51.98	272	958	56.31	7.30	1.19
publication	974	4,002	242.32	42.42	163	1,781	33.75	13.58	1.19
method	3,311	50,067	823.75	530.64	439	9,074	90.89	69.19	1.19
build	1,414	12,739	351.79	135.02	340	4,421	70.39	33.71	1.19
code	1,478	14,055	367.71	148.96	216	3,212	44.72	24.49	1.19
participation	1,201	8,721	298.80	92.43	272	2,408	56.31	18.36	1.19
rubric	782	533	194.55	5.65	44	237	9.11	1.81	1.19
base	3,674	57,720	914.06	611.75	456	10,808	94.41	82.41	1.19
curriculum	935	3,611	232.62	38.27	141	632	29.19	4.82	1.19
workshop	830	1,530	206.50	16.22	130	478	26.92	3.64	1.19
measure	3,805	60,920	946.65	645.67	377	8,666	78.05	66.08	1.18
test	3,472	54,451	863.80	577.10	327	7,484	67.70	57.06	1.18
perspective	1,416	13,789	352.29	146.14	341	4,144	70.60	31.60	1.18
care	1,369	12,929	340.59	137.03	184	3,187	38.10	24.30	1.18

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
topic	1,062	6,818	264.22	72.26	222	2,404	45.96	18.33	1.18
factor	3,221	49,943	801.36	529.32	359	8,476	74.33	64.63	1.18
indigenous	896	3,617	222.92	38.34	56	617	11.59	4.70	1.18
teach	990	5,671	246.30	60.10	164	1,662	33.95	12.67	1.18
instrument	1,107	8,079	275.41	85.63	181	2,401	37.47	18.31	1.18
grant	1,016	6,339	252.77	67.18	195	2,397	40.37	18.28	1.17
rater	728	738	181.12	7.82	24	148	4.97	1.13	1.17
bias	1,161	9,426	288.85	99.90	215	2,573	44.51	19.62	1.17
department	1,016	6,586	252.77	69.80	241	2,054	49.90	15.66	1.17
guide	931	5,117	231.62	54.23	291	2,585	60.25	19.71	1.17
lesson	832	3,164	206.99	33.53	191	1,082	39.54	8.25	1.17
graduate	804	2,580	200.03	27.34	135	787	27.95	6.00	1.17
client	1,045	7,537	259.99	79.88	107	876	22.15	6.68	1.17
effectiveness	916	4,885	227.89	51.77	254	1,813	52.59	13.82	1.17
competency	729	1,244	181.37	13.18	117	341	24.22	2.60	1.17
system	4,674	80,666	1,162.85	854.95	364	9,277	75.36	70.74	1.17
motivation	975	6,284	242.57	66.60	192	2,044	39.75	15.59	1.17
understanding	1,477	16,536	367.46	175.26	347	5,323	71.84	40.59	1.16
math	688	549	171.17	5.82	50	156	10.35	1.19	1.16
content	1,661	20,206	413.24	214.15	310	4,499	64.18	34.30	1.16
communication	1,216	11,308	302.53	119.85	230	2,826	47.62	21.55	1.16
logic	1,019	7,363	253.52	78.04	152	1,668	31.47	12.72	1.16
aim	1,262	12,439	313.97	131.84	344	5,277	71.22	40.24	1.16
inform	901	5,225	224.16	55.38	312	2,542	64.60	19.38	1.16
influence	2,200	31,765	547.34	336.66	354	7,451	73.29	56.81	1.16
access	1,320	13,862	328.40	146.92	281	3,915	58.18	29.85	1.16
scale	1,808	24,116	449.81	255.60	265	4,983	54.87	37.99	1.16
leadership	889	5,409	221.18	57.33	178	1,064	36.85	8.11	1.16
purpose	1,411	16,065	351.04	170.27	371	6,302	76.81	48.05	1.15
principal	910	5,997	226.40	63.56	114	2,185	23.60	16.66	1.15
ensure	1,128	10,565	280.64	111.97	332	4,663	68.74	35.55	1.15
organizational	982	7,709	244.31	81.70	155	1,016	32.09	7.75	1.15
pre	618	275	153.75	2.91	122	112	25.26	0.85	1.15
multiple	1,353	15,415	336.61	163.38	352	5,417	72.88	41.30	1.15
respondent	1,053	9,586	261.98	101.60	157	1,153	32.51	8.79	1.15
estimate	2,054	30,027	511.02	318.24	175	5,026	36.23	38.32	1.15
facilitate	959	7,691	238.59	81.51	290	3,703	60.04	28.23	1.15
expert	866	5,801	215.45	61.48	201	1,477	41.61	11.26	1.15

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
what	4,269	75,784	1,062.09	803.20	431	8,157	89.23	62.20	1.14
among	2,501	39,497	622.23	418.61	399	8,576	82.61	65.39	1.14
you	2,040	29,939	507.53	317.31	217	3,306	44.93	25.21	1.14
competence	749	3,528	186.34	37.39	110	888	22.77	6.77	1.14
institution	1,212	13,114	301.53	138.99	211	2,557	43.69	19.50	1.14
actor	874	6,521	217.44	69.11	110	1,271	22.77	9.69	1.14
collaborative	639	1,764	158.98	18.70	178	525	36.85	4.00	1.14
evaluative	612	1,209	152.26	12.81	121	282	25.05	2.15	1.14
agency	1,021	9,629	254.02	102.05	220	1,993	45.55	15.20	1.14
innovation	885	6,845	220.18	72.55	152	1,320	31.47	10.06	1.14
performance	2,533	40,989	630.19	434.43	279	4,785	57.76	36.48	1.14
understand	1,751	24,874	435.63	263.63	391	7,136	80.95	54.41	1.14
reliability	803	5,314	199.78	56.32	153	1,407	31.68	10.73	1.14
causal	776	4,928	193.06	52.23	124	1,113	25.67	8.49	1.13
sector	1,064	10,901	264.71	115.54	192	1,664	39.75	12.69	1.13
monitoring	679	2,969	168.93	31.47	129	1,054	26.71	8.04	1.13
accountability	621	1,770	154.50	18.76	157	481	32.51	3.67	1.13
meeting	791	5,387	196.79	57.09	178	1,674	36.85	12.76	1.13
participatory	566	748	140.82	7.93	120	243	24.84	1.85	1.13
relate	1,980	30,064	492.61	318.64	393	8,464	81.37	64.54	1.13
network	1,772	25,714	440.86	272.53	189	3,147	39.13	24.00	1.13
sd	692	3,408	172.16	36.12	113	753	23.40	5.74	1.13
effective	1,202	14,045	299.05	148.86	329	5,045	68.12	38.47	1.13
success	1,006	9,915	250.28	105.08	278	3,452	57.56	26.32	1.13
quantitative	768	4,973	191.07	52.71	209	2,130	43.27	16.24	1.13
mathematics	601	1,584	149.52	16.79	65	434	13.46	3.31	1.13
critical	1,230	14,773	306.01	156.57	300	4,866	62.11	37.10	1.13
promote	1,018	10,428	253.27	110.52	282	3,759	58.39	28.66	1.13
attend	757	4,927	188.33	52.22	205	1,942	42.44	14.81	1.13
complete	1,435	19,131	357.01	202.76	333	7,015	68.94	53.49	1.13
component	1,644	23,569	409.01	249.80	282	5,807	58.39	44.28	1.13
instruction	813	6,314	202.27	66.92	138	1,972	28.57	15.04	1.13
organisation	718	4,487	178.63	47.56	96	854	19.88	6.51	1.13
thinking	673	3,545	167.44	37.57	175	1,431	36.23	10.91	1.13
parent	1,131	13,149	281.38	139.36	137	1,992	28.36	15.19	1.13
explore	1,011	10,596	251.53	112.30	316	4,586	65.42	34.97	1.13
highlight	830	6,839	206.50	72.48	304	3,411	62.94	26.01	1.13
foundation	708	4,331	176.14	45.90	226	2,160	46.79	16.47	1.13

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
business	1,211	14,923	301.29	158.16	154	2,549	31.88	19.44	1.12
recommendation	709	4,407	176.39	46.71	191	1,283	39.54	9.78	1.12
positive	1,901	29,384	472.95	311.43	337	6,561	69.77	50.03	1.12
characteristic	1,570	22,449	390.60	237.93	276	6,754	57.14	51.50	1.12
objective	1,039	11,420	258.49	121.04	270	3,943	55.90	30.06	1.12
collection	961	9,739	239.09	103.22	285	3,345	59.01	25.51	1.12
specific	2,002	31,608	498.08	335.00	401	8,469	83.02	64.57	1.12
national	1,577	22,705	392.34	240.64	303	4,057	62.73	30.93	1.12
session	813	6,793	202.27	72.00	144	1,253	29.81	9.55	1.12
setting	985	10,413	245.06	110.36	276	3,481	57.14	26.54	1.12
meet	983	10,416	244.56	110.39	305	4,364	63.15	33.27	1.12
practitioner	627	2,937	155.99	31.13	170	1,079	35.20	8.23	1.12
delivery	684	4,260	170.17	45.15	163	1,241	33.75	9.46	1.12
metric	642	3,460	159.72	36.67	82	683	16.98	5.21	1.12
contextual	597	2,453	148.53	26.00	163	828	33.75	6.31	1.12
who	4,389	82,004	1,091.94	869.13	453	8,011	93.79	61.08	1.12
ask	1,290	17,173	320.94	182.01	334	4,865	69.15	37.09	1.12
effort	1,245	16,222	309.74	171.93	342	4,860	70.81	37.06	1.12
chapter	603	2,897	150.02	30.70	92	1,079	19.05	8.23	1.12
perceive	1,131	14,152	281.38	149.99	238	3,434	49.28	26.18	1.11
our	4,934	94,431	1,227.53	1,000.83	396	10,352	81.99	78.93	1.11
literacy	553	2,062	137.58	21.85	81	361	16.77	2.75	1.11
role	2,429	41,705	604.31	442.01	380	8,503	78.67	64.83	1.11
relevant	1,191	15,517	296.31	164.46	337	5,305	69.77	40.45	1.11
contribute	1,130	14,276	281.13	151.31	354	5,976	73.29	45.57	1.11
consultant	503	1,054	125.14	11.17	65	377	13.46	2.87	1.11
collect	1,000	11,612	248.79	123.07	294	4,686	60.87	35.73	1.11
fund	857	8,627	213.21	91.43	240	1,574	49.69	12.00	1.11
scientific	778	6,944	193.56	73.60	155	1,904	32.09	14.52	1.11
they	8,420	168,505	2,094.82	1,785.91	470	11,820	97.31	90.13	1.11
through	3,507	64,627	872.51	684.95	450	10,608	93.17	80.88	1.11
external	933	10,283	232.12	108.99	240	3,190	49.69	24.32	1.11
limitation	791	7,278	196.79	77.14	306	3,557	63.35	27.12	1.11
toc	455	355	113.20	3.76	17	41	3.52	0.31	1.11
overall	1,231	16,745	306.26	177.47	322	5,863	66.67	44.70	1.11
perception	986	11,571	245.31	122.64	208	2,718	43.06	20.72	1.11
adapt	672	4,985	167.19	52.83	239	2,564	49.48	19.55	1.11
standardized	550	2,485	136.83	26.34	126	1,080	26.09	8.23	1.11

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
internal	1,028	12,591	255.76	133.45	226	4,065	46.79	31.00	1.11
food	845	8,739	210.23	92.62	89	1,947	18.43	14.85	1.11
cjpe	432	0	107.48	0.00	30	0	6.21	0.00	1.11
document	965	11,309	240.08	119.86	267	3,638	55.28	27.74	1.11
methodology	675	5,173	167.93	54.83	226	2,298	46.79	17.52	1.11
reading	916	10,523	227.89	111.53	101	2,303	20.91	17.56	1.11
decision	1,657	26,216	412.25	277.85	332	4,392	68.74	33.49	1.11
instructional	481	1,300	119.67	13.78	74	271	15.32	2.07	1.10
patton	424	101	105.49	1.07	99	53	20.50	0.40	1.10
theme	734	6,790	182.61	71.96	163	2,026	33.75	15.45	1.10
faculty	568	3,307	141.31	35.05	104	623	21.53	4.75	1.10
guideline	562	3,136	139.82	33.24	207	1,250	42.86	9.53	1.10
mentor	459	943	114.19	9.99	39	186	8.07	1.42	1.10
information	3,002	55,182	746.87	584.85	420	7,626	86.96	58.15	1.10
implementer	408	35	101.51	0.37	43	29	8.90	0.22	1.10
individual	3,050	56,329	758.81	597.01	392	8,247	81.16	62.88	1.10
citation	443	866	110.21	9.18	47	303	9.73	2.31	1.10
future	1,409	21,518	350.55	228.06	374	6,198	77.43	47.26	1.10
reflection	643	5,101	159.97	54.06	170	1,994	35.20	15.20	1.10
discipline	634	4,924	157.73	52.19	147	1,552	30.43	11.83	1.10
charter	474	1,592	117.93	16.87	20	300	4.14	2.29	1.10
within	3,396	64,063	844.89	678.98	431	10,363	89.23	79.02	1.10
regard	1,561	24,860	388.36	263.48	341	7,170	70.60	54.67	1.10
offer	1,409	21,554	350.55	228.44	352	5,958	72.88	45.43	1.10
informal	526	2,742	130.86	29.06	100	1,159	20.70	8.84	1.10
your	775	8,181	192.81	86.71	154	2,121	31.88	16.17	1.10
stem	630	5,028	156.74	53.29	113	1,900	23.40	14.49	1.10
realist	425	776	105.74	8.22	32	195	6.63	1.49	1.10
enrollment	423	667	105.24	7.07	48	175	9.94	1.33	1.10
facilitator	412	479	102.50	5.08	90	190	18.63	1.45	1.10
discuss	1,423	22,078	354.03	234.00	360	7,029	74.53	53.60	1.10
administrative	540	3,312	134.35	35.10	117	1,048	24.22	7.99	1.10
align	516	2,757	128.38	29.22	208	1,367	43.06	10.42	1.10
enroll	438	1,086	108.97	11.51	85	490	17.60	3.74	1.10
aspect	1,204	17,556	299.54	186.07	315	5,896	65.22	44.96	1.10
value	4,664	91,854	1,160.36	973.52	420	9,941	86.96	75.80	1.10
sustainable	472	1,930	117.43	20.46	113	573	23.40	4.37	1.10
pilot	454	1,537	112.95	16.29	131	654	27.12	4.99	1.10

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
promotion	505	2,685	125.64	28.46	135	979	27.95	7.46	1.09
applicant	441	1,433	109.72	15.19	41	230	8.49	1.75	1.09
broad	881	10,911	219.18	115.64	287	4,566	59.42	34.82	1.09
lack	1,272	19,403	316.46	205.64	342	7,055	70.81	53.79	1.09
technical	637	5,705	158.48	60.46	189	2,283	39.13	17.41	1.09
priority	571	4,328	142.06	45.87	185	1,544	38.30	11.77	1.09
effect	4,742	94,205	1,179.77	998.44	355	10,124	73.50	77.19	1.09
treatment	1,621	27,013	403.29	286.30	145	4,581	30.02	34.93	1.09
housing	614	5,452	152.76	57.78	43	766	8.90	5.84	1.09
and	109,621	2,354,038	27,272.70	24,949.46	483	13,115	100.00	100.00	1.09
discussion	1,422	22,869	353.78	242.38	407	9,388	84.27	71.58	1.09
available	1,382	21,991	343.83	233.07	383	7,509	79.30	57.26	1.09
barrier	646	6,094	160.72	64.59	171	1,705	35.40	13.00	1.09
foster	530	3,750	131.86	39.74	165	1,565	34.16	11.93	1.09
progress	643	6,156	159.97	65.24	218	2,892	45.13	22.05	1.09
diverse	579	4,796	144.05	50.83	220	2,498	45.55	19.05	1.09
systematic	563	4,431	140.07	46.96	214	2,308	44.31	17.60	1.09
rating	637	6,148	158.48	65.16	91	855	18.84	6.52	1.09
importance	1,090	15,873	271.18	168.23	341	6,118	70.60	46.65	1.09
feel	913	12,106	227.15	128.31	234	3,500	48.45	26.69	1.09
fit	887	11,539	220.68	122.30	221	3,963	45.76	30.22	1.09
share	1,417	23,005	352.54	243.82	320	5,607	66.25	42.75	1.09
achieve	1,188	18,090	295.56	191.73	330	6,529	68.32	49.78	1.09
useful	833	10,465	207.24	110.91	290	5,111	60.04	38.97	1.09
leader	743	8,489	184.85	89.97	176	1,688	36.44	12.87	1.09
provider	518	3,688	128.87	39.09	106	720	21.95	5.49	1.09
formative	384	788	95.54	8.35	106	274	21.95	2.09	1.09
step	1,359	21,938	338.11	232.51	300	6,033	62.11	46.00	1.09
profession	475	2,761	118.18	29.26	110	756	22.77	5.76	1.09
ranking	436	2,069	108.47	21.93	42	585	8.70	4.46	1.09
create	1,537	25,882	382.39	274.31	361	6,655	74.74	50.74	1.09
reflect	1,318	21,097	327.91	223.60	365	7,037	75.57	53.66	1.09
primary	1,031	14,904	256.50	157.96	288	5,286	59.63	40.30	1.09
learner	589	5,486	146.54	58.14	87	450	18.01	3.43	1.08
el	420	1,794	104.49	19.01	47	370	9.73	2.82	1.08
think	1,473	24,547	366.47	260.16	294	5,987	60.87	45.65	1.08
educator	410	1,638	102.00	17.36	89	419	18.43	3.19	1.08

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
testing	611	6,124	152.01	64.91	145	2,069	30.02	15.78	1.08
ongoing	480	3,306	119.42	35.04	196	1,894	40.58	14.44	1.08
consulting	351	434	87.33	4.60	43	202	8.90	1.54	1.08
yes	408	1,745	101.51	18.49	96	816	19.88	6.22	1.08
issue	2,057	37,617	511.76	398.69	410	7,036	84.89	53.65	1.08
institutional	725	8,639	180.37	91.56	176	1,989	36.44	15.17	1.08
post	523	4,301	130.12	45.58	150	1,808	31.06	13.79	1.08
governance	527	4,503	131.11	47.73	84	686	17.39	5.23	1.08
baseline	524	4,427	130.37	46.92	99	1,129	20.50	8.61	1.08
canada	460	3,006	114.44	31.86	95	886	19.67	6.76	1.08
degree	1,263	20,504	314.22	217.31	291	6,522	60.25	49.73	1.08
contribution	838	11,236	208.49	119.09	234	4,294	48.45	32.74	1.08
enable	777	9,897	193.31	104.89	236	4,165	48.86	31.76	1.08
prevention	409	1,955	101.76	20.72	93	709	19.25	5.41	1.08
beneficiary	348	681	86.58	7.22	71	267	14.70	2.04	1.08
environment	1,311	21,645	326.16	229.41	299	5,076	61.90	38.70	1.08
select	918	13,059	228.39	138.41	297	5,122	61.49	39.05	1.08
insight	600	6,176	149.27	65.46	250	3,329	51.76	25.38	1.08
girl	543	4,899	135.09	51.92	40	883	8.28	6.73	1.08
deliver	517	4,382	128.62	46.44	180	2,122	37.27	16.18	1.08
center	820	11,039	204.01	117.00	202	3,636	41.82	27.72	1.08
shared	384	1,621	95.54	17.18	112	785	23.19	5.99	1.08
important	2,679	51,642	666.51	547.33	444	10,751	91.93	81.97	1.08
ability	1,248	20,466	310.49	216.91	315	6,030	65.22	45.98	1.08
methodological	413	2,324	102.75	24.63	166	1,100	34.37	8.39	1.08
institute	448	3,230	111.46	34.23	187	1,871	38.72	14.27	1.08
potential	1,620	28,801	403.04	305.25	376	7,690	77.85	58.64	1.08
additional	1,075	16,887	267.45	178.98	318	6,629	65.84	50.55	1.08
list	910	13,281	226.40	140.76	302	4,527	62.53	34.52	1.08
encourage	637	7,316	158.48	77.54	265	3,200	54.87	24.40	1.08
demographic	454	3,449	112.95	36.55	145	1,162	30.02	8.86	1.07
feasibility	333	787	82.85	8.34	68	464	14.08	3.54	1.07
enhance	706	8,893	175.65	94.25	255	3,896	52.80	29.71	1.07
practical	652	7,735	162.21	81.98	252	3,240	52.17	24.70	1.07
comprehensive	472	3,784	117.43	40.11	210	2,155	43.48	16.43	1.07
pupil	443	3,161	110.21	33.50	25	435	5.18	3.32	1.07
non	344	1,014	85.58	10.75	123	531	25.47	4.05	1.07
aea	295	46	73.39	0.49	29	10	6.00	0.08	1.07

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
prior	1,003	15,563	249.54	164.95	250	4,935	51.76	37.63	1.07
strengthen	415	2,692	103.25	28.53	173	1,673	35.82	12.76	1.07
email	328	721	81.60	7.64	201	174	41.61	1.33	1.07
writing	533	5,343	132.61	56.63	86	1,631	17.81	12.44	1.07
reflective	354	1,438	88.07	15.24	83	606	17.18	4.62	1.07
studies	339	1,050	84.34	11.13	115	540	23.81	4.12	1.07
often	1,826	33,693	454.29	357.10	384	8,333	79.50	63.54	1.07
variance	763	10,388	189.83	110.10	118	2,396	24.43	18.27	1.07
screening	393	2,399	97.77	25.43	43	764	8.90	5.83	1.07
ise	293	143	72.90	1.52	8	17	1.66	0.13	1.07
expertise	424	3,081	105.49	32.65	180	1,035	37.27	7.89	1.07
efficacy	417	2,970	103.75	31.48	109	1,119	22.57	8.53	1.07
cc	314	740	78.12	7.84	20	172	4.14	1.31	1.07
grantee	283	23	70.41	0.24	20	11	4.14	0.08	1.07
dimension	898	13,501	223.41	143.09	166	3,657	34.37	27.88	1.07
gap	592	6,798	147.28	72.05	212	2,371	43.89	18.08	1.07
decision-making	410	2,935	102.00	31.11	136	1,003	28.16	7.65	1.07
re	345	1,481	85.83	15.70	109	341	22.57	2.60	1.07
principle	1,145	19,076	284.87	202.18	229	4,265	47.41	32.52	1.07
specifically	798	11,399	198.54	120.81	296	5,149	61.28	39.26	1.07
gather	405	2,777	100.76	29.43	161	1,690	33.33	12.89	1.07
pa	371	2,168	92.30	22.98	40	522	8.28	3.98	1.07
ev	304	692	75.63	7.33	43	152	8.90	1.16	1.07
new	3,668	74,689	912.56	791.60	435	9,629	90.06	73.42	1.07
figure	2,263	43,731	563.01	463.49	319	6,651	66.05	50.71	1.07
nurse	324	1,119	80.61	11.86	38	362	7.87	2.76	1.07
descriptive	404	2,974	100.51	31.52	160	1,399	33.13	10.67	1.07
ea	295	524	73.39	5.55	35	129	7.25	0.98	1.07
career	469	4,417	116.68	46.81	121	1,210	25.05	9.23	1.07
fidelity	316	1,124	78.62	11.91	60	246	12.42	1.88	1.07
earnings	266	12	66.18	0.13	20	9	4.14	0.07	1.07
interest	1,655	30,597	411.75	324.28	376	7,016	77.85	53.50	1.07
mechanism	1,285	22,401	319.70	237.42	215	5,291	44.51	40.34	1.07
publish	665	8,830	165.45	93.59	244	3,597	50.52	27.43	1.07
trust	573	6,757	142.56	71.61	146	1,435	30.23	10.94	1.07
van	483	4,786	120.17	50.72	154	1,568	31.88	11.96	1.07
miss	447	4,003	111.21	42.43	161	2,227	33.33	16.98	1.07
limited	621	7,968	154.50	84.45	263	4,388	54.45	33.46	1.07

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
instructor	329	1,481	81.85	15.70	48	255	9.94	1.94	1.07
serve	949	15,126	236.10	160.31	298	5,701	61.70	43.47	1.07
ecb	266	234	66.18	2.48	9	20	1.86	0.15	1.06
apply	1,600	29,608	398.07	313.80	373	7,989	77.23	60.91	1.06
application	1,131	19,290	281.38	204.45	260	5,159	53.83	39.34	1.06
respond	693	9,610	172.41	101.85	253	3,960	52.38	30.19	1.06
integrate	510	5,606	126.88	59.42	198	2,695	40.99	20.55	1.06
empowerment	288	808	71.65	8.56	75	271	15.53	2.07	1.06
evidence-based	271	351	67.42	3.72	88	121	18.22	0.92	1.06
remedial	263	215	65.43	2.28	7	121	1.45	0.92	1.06
more	8,411	180,263	2,092.58	1,910.53	474	12,625	98.14	96.26	1.06
area	2,191	42,749	545.10	453.08	361	7,349	74.74	56.04	1.06
maker	350	2,153	87.08	22.82	122	834	25.26	6.36	1.06
centre	602	7,783	149.77	82.49	119	2,213	24.64	16.87	1.06
model	5,641	119,143	1,403.43	1,262.75	388	9,100	80.33	69.39	1.06
manage	551	6,698	137.08	70.99	210	2,709	43.48	20.66	1.06
intend	546	6,546	135.84	69.38	239	3,045	49.48	23.22	1.06
revise	360	2,420	89.56	25.65	208	1,333	43.06	10.16	1.06
five	953	15,652	237.10	165.89	312	5,969	64.60	45.51	1.06
examine	1,499	27,780	372.94	294.43	323	7,177	66.87	54.72	1.06
mental	565	7,070	140.57	74.93	108	1,505	22.36	11.48	1.06
strategic	495	5,562	123.15	58.95	141	1,432	29.19	10.92	1.06
assignment	417	3,807	103.75	40.35	89	1,129	18.43	8.61	1.06
assistance	376	2,922	93.55	30.97	130	1,200	26.92	9.15	1.06
attendance	314	1,513	78.12	16.04	78	413	16.15	3.15	1.06
people	2,248	44,443	559.28	471.03	320	5,179	66.25	39.49	1.06
read	793	12,180	197.29	129.09	197	3,432	40.79	26.17	1.06
reform	652	9,086	162.21	96.30	94	1,469	19.46	11.20	1.06
monitor	512	5,962	127.38	63.19	165	2,617	34.16	19.95	1.06
hospital	479	5,307	119.17	56.25	63	972	13.04	7.41	1.06
wam	238	22	59.21	0.23	8	3	1.66	0.02	1.06
response	2,279	45,317	566.99	480.30	307	7,151	63.56	54.53	1.06
note	2,089	41,060	519.72	435.18	395	9,120	81.78	69.54	1.06
analyze	827	13,066	205.75	138.48	259	4,623	53.62	35.25	1.06
panel	657	9,287	163.46	98.43	86	1,750	17.81	13.34	1.06
recruitment	344	2,456	85.58	26.03	81	774	16.77	5.90	1.06
guidance	344	2,443	85.58	25.89	156	1,021	32.30	7.78	1.06
na	295	1,344	73.39	14.24	35	437	7.25	3.33	1.06

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
holistic	269	812	66.92	8.61	73	394	15.11	3.00	1.06
hf	260	623	64.69	6.60	6	133	1.24	1.01	1.06
mindset	243	256	60.46	2.71	32	138	6.63	1.05	1.06
funder	240	133	59.71	1.41	110	53	22.77	0.40	1.06
high	4,511	94,878	1,122.30	1,005.57	412	11,229	85.30	85.62	1.06
well	3,505	72,566	872.01	769.10	468	11,973	96.89	91.29	1.06
statistic	527	6,504	131.11	68.93	163	2,140	33.75	16.32	1.06
appendix	404	3,750	100.51	39.74	104	1,382	21.53	10.54	1.06
recommend	386	3,412	96.03	36.16	192	1,947	39.75	14.85	1.06
inquiry	376	3,162	93.55	33.51	105	1,021	21.74	7.78	1.06
my	1,311	23,986	326.16	254.22	191	3,345	39.54	25.51	1.06
interdisciplinary	248	401	61.70	4.25	47	215	9.73	1.64	1.06
aej	230	0	57.22	0.00	47	0	9.73	0.00	1.06
action	1,589	30,196	395.33	320.03	294	5,235	60.87	39.92	1.06
concept	1,234	22,353	307.01	236.91	289	4,681	59.83	35.69	1.06
measurement	1,000	17,084	248.79	181.07	217	3,719	44.93	28.36	1.06
agree	630	8,936	156.74	94.71	235	3,856	48.65	29.40	1.06
awareness	428	4,415	106.48	46.79	171	1,727	35.40	13.17	1.06
meaningful	368	3,110	91.56	32.96	157	1,796	32.51	13.69	1.06
campus	267	805	66.43	8.53	48	267	9.94	2.04	1.06
me	813	13,109	202.27	138.94	220	2,749	45.55	20.96	1.06
additionally	395	3,789	98.27	40.16	155	2,052	32.09	15.65	1.06
innovative	319	2,054	79.36	21.77	123	902	25.47	6.88	1.06
sciences	253	642	62.94	6.80	109	421	22.57	3.21	1.06
wow	225	38	55.98	0.40	16	33	3.31	0.25	1.06
construct	1,148	20,493	285.61	217.20	231	5,507	47.83	41.99	1.06
incentive	490	5,932	121.91	62.87	116	1,354	24.02	10.32	1.06
round	394	3,744	98.02	39.68	73	1,450	15.11	11.06	1.06
video	364	3,096	90.56	32.81	65	720	13.46	5.49	1.06
budget	363	3,088	90.31	32.73	117	833	24.22	6.35	1.06
county	359	3,002	89.32	31.82	41	546	8.49	4.16	1.06
conference	356	2,878	88.57	30.50	132	942	27.33	7.18	1.06
hire	310	1,932	77.13	20.48	83	674	17.18	5.14	1.06
covariate	248	569	61.70	6.03	43	175	8.90	1.33	1.06
toward	743	11,641	184.85	123.38	223	3,744	46.17	28.55	1.06
background	680	10,245	169.18	108.58	248	3,999	51.35	30.49	1.06
collective	541	7,137	134.60	75.64	114	1,671	23.60	12.74	1.06
confidence	494	6,056	122.90	64.19	164	2,108	33.95	16.07	1.06

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
acknowledge	453	5,119	112.70	54.25	217	2,680	44.93	20.43	1.06
conceptual	452	5,111	112.45	54.17	155	1,786	32.09	13.62	1.06
conversation	363	3,159	90.31	33.48	116	1,077	24.02	8.21	1.06
three	2,680	54,875	666.76	581.60	424	10,746	87.78	81.94	1.05
intended	253	814	62.94	8.63	107	498	22.15	3.80	1.05
dissemination	252	787	62.70	8.34	102	451	21.12	3.44	1.05
ethical	422	4,544	104.99	48.16	124	1,010	25.67	7.70	1.05
admission	285	1,515	70.91	16.06	31	579	6.42	4.41	1.05
completion	319	2,405	79.36	25.49	116	1,176	24.02	8.97	1.05
coordinator	231	374	57.47	3.96	52	141	10.77	1.08	1.05
involvement	496	6,283	123.40	66.59	187	2,159	38.72	16.46	1.05
deep	483	6,031	120.17	63.92	160	2,727	33.13	20.79	1.05
rigorous	262	1,181	65.18	12.52	121	860	25.05	6.56	1.05
nations	255	996	63.44	10.56	54	403	11.18	3.07	1.05
subscale	242	691	60.21	7.32	32	121	6.63	0.92	1.05
cognitive	632	9,460	157.24	100.26	135	1,669	27.95	12.73	1.05
statistically	488	6,179	121.41	65.49	122	2,230	25.26	17.00	1.05
embed	435	4,998	108.22	52.97	149	2,213	30.85	16.87	1.05
undertake	416	4,623	103.50	49.00	165	2,316	34.16	17.66	1.05
healthy	371	3,587	92.30	38.02	76	1,376	15.73	10.49	1.05
randomize	270	1,348	67.17	14.29	74	518	15.32	3.95	1.05
st	344	3,118	85.58	33.05	77	717	15.94	5.47	1.05
culturally	276	1,585	68.67	16.80	77	753	15.94	5.74	1.05
eligible	246	959	61.20	10.16	72	471	14.91	3.59	1.05
keywords	227	509	56.48	5.39	227	294	47.00	2.24	1.05
readiness	224	420	55.73	4.45	68	223	14.08	1.70	1.05
variable	2,295	46,643	570.97	494.35	248	5,461	51.35	41.64	1.05
personal	804	13,394	200.03	141.96	223	3,795	46.17	28.94	1.05
capture	632	9,494	157.24	100.62	223	3,640	46.17	27.75	1.05
inclusion	424	4,925	105.49	52.20	187	2,192	38.72	16.71	1.05
ministry	314	2,568	78.12	27.22	85	644	17.60	4.91	1.05
subgroup	294	2,119	73.14	22.46	48	573	9.94	4.37	1.05
organisational	260	1,279	64.69	13.56	51	252	10.56	1.92	1.05
responsive	259	1,306	64.44	13.84	97	724	20.08	5.52	1.05
supervisor	251	1,139	62.45	12.07	47	268	9.73	2.04	1.05
current	1,355	25,843	337.11	273.90	331	6,490	68.53	49.49	1.05
satisfaction	418	4,874	103.99	51.66	90	988	18.63	7.53	1.05
utilize	405	4,579	100.76	48.53	159	2,160	32.92	16.47	1.05

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
tension	397	4,349	98.77	46.09	86	1,783	17.81	13.60	1.05
programming	279	1,844	69.41	19.54	91	552	18.84	4.21	1.05
counterfactual	234	787	58.22	8.34	39	136	8.07	1.04	1.05
sample	2,707	56,129	673.48	594.89	287	6,367	59.42	48.55	1.05
country	1,607	31,577	399.81	334.67	233	3,588	48.24	27.36	1.05
start	1,033	18,676	257.00	197.94	320	6,649	66.25	50.70	1.05
observation	963	17,098	239.59	181.21	202	5,817	41.82	44.35	1.05
expectation	578	8,518	143.80	90.28	226	2,925	46.79	22.30	1.05
complexity	576	8,511	143.30	90.20	194	2,968	40.17	22.63	1.05
valuable	354	3,511	88.07	37.21	186	2,161	38.51	16.48	1.05
semester	225	589	55.98	6.24	39	170	8.07	1.30	1.05
lee	279	1,962	69.41	20.79	109	736	22.57	5.61	1.05
website	234	954	58.22	10.11	85	417	17.60	3.18	1.05
nutrition	229	763	56.97	8.09	51	249	10.56	1.90	1.05
attrition	214	490	53.24	5.19	46	185	9.52	1.41	1.05
pre-service	200	138	49.76	1.46	12	28	2.48	0.21	1.05
answer	777	13,042	193.31	138.23	276	4,124	57.14	31.44	1.05
checklist	210	436	52.25	4.62	60	170	12.42	1.30	1.05
colleague	349	3,616	86.83	38.32	153	1,586	31.68	12.09	1.05
summer	324	3,026	80.61	32.07	53	1,038	10.97	7.91	1.05
analyse	443	5,815	110.21	61.63	114	2,328	23.60	17.75	1.05
administer	338	3,392	84.09	35.95	141	1,518	29.19	11.57	1.05
collaborate	215	644	53.49	6.83	112	431	23.19	3.29	1.05
summative	191	134	47.52	1.42	66	55	13.66	0.42	1.05
population	1,334	25,803	331.89	273.48	276	4,644	57.14	35.41	1.05
independent	937	16,837	233.12	178.45	220	6,064	45.55	46.24	1.05
final	863	15,207	214.71	161.17	277	5,901	57.35	44.99	1.05
output	711	11,834	176.89	125.42	146	2,115	30.23	16.13	1.05
user	684	11,157	170.17	118.25	146	1,443	30.23	11.00	1.05
difficulty	637	10,166	158.48	107.75	214	4,123	44.31	31.44	1.05
canadian	275	2,128	68.42	22.55	76	495	15.73	3.77	1.05
reporting	271	2,004	67.42	21.24	104	582	21.53	4.44	1.05
validation	260	1,767	64.69	18.73	80	739	16.56	5.63	1.05
administrator	239	1,288	59.46	13.65	87	521	18.01	3.97	1.05
healthcare	213	766	52.99	8.12	59	272	12.22	2.07	1.05
caregiver	200	440	49.76	4.66	30	123	6.21	0.94	1.05
community-based	199	423	49.51	4.48	65	167	13.46	1.27	1.05

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
selection	710	11,898	176.64	126.10	182	3,171	37.68	24.18	1.05
core	650	10,523	161.71	111.53	219	3,055	45.34	23.29	1.05
south	585	9,069	145.54	96.12	113	2,008	23.40	15.31	1.05
successful	559	8,540	139.07	90.51	237	3,798	49.07	28.96	1.05
communicate	322	3,157	80.11	33.46	158	1,555	32.71	11.86	1.05
recruit	304	2,806	75.63	29.74	117	1,332	24.22	10.16	1.05
elementary	272	2,010	67.67	21.30	55	754	11.39	5.75	1.05
africa	378	4,569	94.04	48.42	54	864	11.18	6.59	1.04
relevance	360	4,093	89.56	43.38	153	2,065	31.68	15.75	1.04
validate	285	2,430	70.91	25.75	127	1,406	26.29	10.72	1.04
operational	264	2,005	65.68	21.25	90	853	18.63	6.50	1.04
reviewer	210	717	52.25	7.60	78	356	16.15	2.71	1.04
low-income	202	572	50.26	6.06	32	176	6.63	1.34	1.04
nursing	198	507	49.26	5.37	33	187	6.83	1.43	1.04
pawson	177	14	44.04	0.15	35	8	7.25	0.06	1.04
establish	1,096	20,676	272.67	219.14	336	7,125	69.57	54.33	1.04
traditional	753	12,964	187.34	137.40	195	4,008	40.37	30.56	1.04
incorporate	507	7,429	126.14	78.74	229	3,682	47.41	28.07	1.04
platform	291	2,612	72.40	27.68	83	897	17.18	6.84	1.04
existing	405	5,233	100.76	55.46	194	2,727	40.17	20.79	1.04
aspiration	236	1,380	58.71	14.63	45	743	9.32	5.67	1.04
graduation	190	392	47.27	4.15	47	128	9.73	0.98	1.04
efl	185	275	46.03	2.91	9	38	1.86	0.29	1.04
bachelor	184	265	45.78	2.81	48	130	9.94	0.99	1.04
ela	176	60	43.79	0.64	11	9	2.28	0.07	1.04
transdisciplinary	175	51	43.54	0.54	14	15	2.90	0.11	1.04
each	4,642	100,626	1,154.89	1,066.49	454	11,771	94.00	89.75	1.04
involve	1,883	38,482	468.47	407.85	391	9,038	80.95	68.91	1.04
seek	835	14,942	207.74	158.36	286	4,685	59.21	35.72	1.04
appropriate	826	14,690	205.50	155.69	312	6,171	64.60	47.05	1.04
assign	565	8,808	140.57	93.35	166	3,395	34.37	25.89	1.04
secondary	523	7,838	130.12	83.07	136	2,907	28.16	22.17	1.04
diversity	430	5,781	106.98	61.27	150	1,894	31.06	14.44	1.04
credibility	231	1,270	57.47	13.46	90	603	18.63	4.60	1.04
towards	699	11,994	173.90	127.12	183	4,209	37.89	32.09	1.04
reasoning	382	4,796	95.04	50.83	72	1,526	14.91	11.64	1.04
supplementary	243	1,650	60.46	17.49	46	527	9.52	4.02	1.04
thematic	213	968	52.99	10.26	85	379	17.60	2.89	1.04

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
coaching	192	503	47.77	5.33	39	119	8.07	0.91	1.04
mentoring	185	360	46.03	3.82	41	90	8.49	0.69	1.04
bibliometric	170	9	42.29	0.10	24	4	4.97	0.03	1.04
cmo	169	38	42.05	0.40	18	1	3.73	0.01	1.04
or	15,298	340,898	3,806.00	3,613.03	479	13,008	99.17	99.18	1.04
while	2,875	60,914	715.27	645.60	431	10,285	89.23	78.42	1.04
extent	857	15,554	213.21	164.85	288	5,974	59.63	45.55	1.04
gain	789	13,947	196.30	147.82	280	4,851	57.97	36.99	1.04
manager	699	11,978	173.90	126.95	146	1,466	30.23	11.18	1.04
beyond	693	11,802	172.41	125.08	280	5,153	57.97	39.29	1.04
story	529	8,135	131.61	86.22	114	1,963	23.60	14.97	1.04
adaptation	353	4,139	87.82	43.87	112	1,328	23.19	10.13	1.04
dialogue	315	3,429	78.37	36.34	87	985	18.01	7.51	1.04
michigan	181	401	45.03	4.25	34	208	7.04	1.59	1.04
nonresponse	171	120	42.54	1.27	15	55	3.11	0.42	1.04
scriven	166	24	41.30	0.25	45	9	9.32	0.07	1.04
direction	739	12,906	183.86	136.79	207	4,678	42.86	35.67	1.04
visit	407	5,457	101.26	57.84	109	1,933	22.57	14.74	1.04
track	338	3,902	84.09	41.36	132	1,701	27.33	12.97	1.04
award	261	2,155	64.93	22.84	86	652	17.81	4.97	1.04
building	672	11,573	167.19	122.66	199	2,263	41.20	17.26	1.04
consistency	313	3,431	77.87	36.36	135	1,475	27.95	11.25	1.04
suspension	262	2,327	65.18	24.66	8	844	1.66	6.44	1.04
investigator	228	1,532	56.72	16.24	42	712	8.70	5.43	1.04
refine	227	1,531	56.48	16.23	124	1,012	25.67	7.72	1.04
ngo	178	400	44.28	4.24	34	136	7.04	1.04	1.04
outreach	171	273	42.54	2.89	43	102	8.90	0.78	1.04
strength	560	9,022	139.32	95.62	222	3,145	45.96	23.98	1.04
format	310	3,378	77.13	35.80	119	1,191	24.64	9.08	1.04
assist	291	2,919	72.40	30.94	152	1,665	31.47	12.70	1.04
invite	269	2,451	66.92	25.98	140	1,458	28.99	11.12	1.04
coach	234	1,619	58.22	17.16	27	204	5.59	1.56	1.04
transparency	219	1,305	54.49	13.83	75	467	15.53	3.56	1.04
exam	208	1,037	51.75	10.99	46	267	9.52	2.04	1.04
robust	326	3,811	81.11	40.39	132	1,905	27.33	14.53	1.04
chen	204	1,040	50.75	11.02	75	430	15.53	3.28	1.04
cb	172	376	42.79	3.99	7	79	1.45	0.60	1.04
pisa	165	168	41.05	1.78	14	41	2.90	0.31	1.04

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
professionalizatio n	164	134	40.80	1.42	11	47	2.28	0.36	1.04
lead	1,996	41,527	496.59	440.13	406	9,858	84.06	75.17	1.04
want	768	13,847	191.07	146.76	232	4,358	48.03	33.23	1.04
description	639	10,927	158.98	115.81	226	4,149	46.79	31.64	1.04
responsibility	554	8,929	137.83	94.63	198	2,303	40.99	17.56	1.04
developmental	323	3,734	80.36	39.58	89	1,080	18.43	8.23	1.04
senior	281	2,819	69.91	29.88	101	964	20.91	7.35	1.04
textbook	208	1,178	51.75	12.49	41	476	8.49	3.63	1.04
well-being	248	2,197	61.70	23.29	91	713	18.84	5.44	1.04
analytic	239	1,915	59.46	20.30	86	741	17.81	5.65	1.04
systemic	228	1,669	56.72	17.69	68	740	14.08	5.64	1.04
proficiency	209	1,302	52.00	13.80	53	281	10.97	2.14	1.04
pe	200	1,031	49.76	10.93	22	223	4.55	1.70	1.04
keyword	162	164	40.30	1.74	117	68	24.22	0.52	1.04
breastfeed	159	154	39.56	1.63	10	26	2.07	0.20	1.04
cousins	156	35	38.81	0.37	35	19	7.25	0.14	1.04
mayne	153	17	38.06	0.18	24	10	4.97	0.08	1.04
tig	152	9	37.82	0.10	14	3	2.90	0.02	1.04
should	2,807	60,043	698.36	636.37	442	10,162	91.51	77.48	1.04
international	940	17,791	233.86	188.56	215	2,782	44.51	21.21	1.04
really	429	6,294	106.73	66.71	133	2,637	27.54	20.11	1.04
check	425	6,156	105.74	65.24	217	2,705	44.93	20.63	1.04
helpful	249	2,185	61.95	23.16	148	1,527	30.64	11.64	1.04
de	781	14,301	194.31	151.57	188	2,889	38.92	22.03	1.04
cohort	281	2,956	69.91	31.33	54	610	11.18	4.65	1.04
retention	264	2,612	65.68	27.68	57	791	11.80	6.03	1.04
infrastructure	264	2,646	65.68	28.04	90	794	18.63	6.05	1.04
turnover	250	2,284	62.20	24.21	31	619	6.42	4.72	1.04
fairness	207	1,335	51.50	14.15	37	364	7.66	2.78	1.04
attainment	204	1,243	50.75	13.17	57	510	11.80	3.89	1.04
interviewee	201	1,221	50.01	12.94	49	297	10.14	2.26	1.04
targeted	194	1,007	48.27	10.67	97	578	20.08	4.41	1.04
retrospective	189	896	47.02	9.50	31	412	6.42	3.14	1.04
eligibility	172	523	42.79	5.54	52	189	10.77	1.44	1.04
psychometric	170	471	42.29	4.99	47	187	9.73	1.43	1.04
homepage	150	33	37.32	0.35	150	23	31.06	0.18	1.04
job	676	11,961	168.18	126.77	166	2,016	34.37	15.37	1.04

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
justice	498	7,879	123.90	83.51	83	1,401	17.18	10.68	1.04
committee	417	6,092	103.75	64.57	143	1,743	29.61	13.29	1.04
stake	224	1,715	55.73	18.18	92	944	19.05	7.20	1.04
american	767	14,083	190.82	149.26	198	2,707	40.99	20.64	1.04
mixed	344	4,506	85.58	47.76	137	1,996	28.36	15.22	1.04
socioeconomic	196	1,152	48.76	12.21	70	467	14.49	3.56	1.04
services	183	912	45.53	9.67	84	481	17.39	3.67	1.04
cronbach	173	643	43.04	6.81	68	270	14.08	2.06	1.04
preschool	169	537	42.05	5.69	22	96	4.55	0.73	1.04
sti	146	54	36.32	0.57	8	17	1.66	0.13	1.04
mentee	143	0	35.58	0.00	7	0	1.45	0.00	1.04
relationship	2,013	42,365	500.82	449.01	355	7,270	73.50	55.43	1.04
scenario	454	6,992	112.95	74.11	78	1,886	16.15	14.38	1.04
annual	395	5,714	98.27	60.56	128	1,742	26.50	13.28	1.04
donor	264	2,716	65.68	28.79	49	654	10.14	4.99	1.04
six	611	10,703	152.01	113.44	230	4,640	47.62	35.38	1.04
supportive	200	1,316	49.76	13.95	93	727	19.25	5.54	1.04
campbell	186	1,066	46.28	11.30	61	372	12.63	2.84	1.04
exploratory	186	1,010	46.28	10.70	81	467	16.77	3.56	1.04
acceptability	166	573	41.30	6.07	24	222	4.97	1.69	1.04
accepted	161	442	40.06	4.68	156	375	32.30	2.86	1.04
condom	159	418	39.56	4.43	8	48	1.66	0.37	1.04
cfa	158	432	39.31	4.58	21	87	4.35	0.66	1.04
um	152	279	37.82	2.96	58	102	12.01	0.78	1.04
pharmacist	145	83	36.07	0.88	10	43	2.07	0.33	1.04
irr	144	92	35.83	0.98	5	13	1.04	0.10	1.04
on	20,903	471,147	5,200.47	4,993.49	482	13,106	99.79	99.93	1.04
implication	611	10,670	152.01	113.09	248	4,583	51.35	34.94	1.04
spend	376	5,310	93.55	56.28	159	2,320	32.92	17.69	1.04
emotional	369	5,171	91.80	54.81	93	1,293	19.25	9.86	1.04
safety	303	3,724	75.38	39.47	72	1,160	14.91	8.84	1.04
volunteer	231	2,065	57.47	21.89	80	659	16.56	5.02	1.04
undergraduate	198	1,279	49.26	13.56	57	464	11.80	3.54	1.04
lab	177	785	44.04	8.32	26	401	5.38	3.06	1.04
societal	193	1,295	48.02	13.73	62	598	12.84	4.56	1.03
obesity	184	1,034	45.78	10.96	26	163	5.38	1.24	1.03
generalizability	159	476	39.56	5.04	54	294	11.18	2.24	1.03
msc	141	93	35.08	0.99	10	28	2.07	0.21	1.03

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
percentage	620	10,951	154.25	116.07	151	3,227	31.26	24.61	1.03
comment	453	7,226	112.70	76.59	173	2,928	35.82	22.33	1.03
train	330	4,358	82.10	46.19	150	1,649	31.06	12.57	1.03
adequate	309	3,878	76.88	41.10	152	2,286	31.47	17.43	1.03
draft	244	2,447	60.70	25.93	101	835	20.91	6.37	1.03
structured	196	1,425	48.76	15.10	91	673	18.84	5.13	1.03
consortium	150	385	37.32	4.08	31	171	6.42	1.30	1.03
maori	149	346	37.07	3.67	8	39	1.66	0.30	1.03
weiss	146	289	36.32	3.06	60	146	12.42	1.11	1.03
app	140	209	34.83	2.22	19	53	3.93	0.40	1.03
heis	138	126	34.33	1.34	7	17	1.45	0.13	1.03
ebp	135	87	33.59	0.92	5	15	1.04	0.11	1.03
whether	1,920	40,635	477.68	430.67	394	8,905	81.57	67.90	1.03
expand	385	5,745	95.78	60.89	191	3,003	39.54	22.90	1.03
scholar	384	5,738	95.54	60.81	119	1,905	24.64	14.53	1.03
federal	350	4,953	87.08	52.49	98	932	20.29	7.11	1.03
united	665	12,231	165.45	129.63	240	2,560	49.69	19.52	1.03
hiv	220	2,086	54.73	22.11	26	202	5.38	1.54	1.03
inspection	209	1,796	52.00	19.04	37	973	7.66	7.42	1.03
interpersonal	208	1,835	51.75	19.45	70	566	14.49	4.32	1.03
inclusive	190	1,372	47.27	14.54	78	652	16.15	4.97	1.03
disciplinary	184	1,291	45.78	13.68	53	397	10.97	3.03	1.03
iterative	169	902	42.05	9.56	84	476	17.39	3.63	1.03
policymaker	153	583	38.06	6.18	64	240	13.25	1.83	1.03
rogers	145	391	36.07	4.14	58	203	12.01	1.55	1.03
schools	144	305	35.83	3.23	33	139	6.83	1.06	1.03
ppp	141	297	35.08	3.15	5	47	1.04	0.36	1.03
meta	138	173	34.33	1.83	37	84	7.66	0.64	1.03
rct	137	184	34.08	1.95	30	24	6.21	0.18	1.03
dif	136	187	33.84	1.98	10	33	2.07	0.25	1.03
elsevier	133	107	33.09	1.13	131	39	27.12	0.30	1.03
australasia	131	45	32.59	0.48	29	34	6.00	0.26	1.03
standard	1,550	32,312	385.63	342.46	303	7,832	62.73	59.72	1.03
able	941	18,518	234.11	196.26	336	6,880	69.57	52.46	1.03
home	659	12,037	163.95	127.58	193	2,928	39.96	22.33	1.03
emerge	658	12,041	163.70	127.62	252	4,508	52.17	34.37	1.03
clinical	558	9,801	138.83	103.88	77	1,853	15.94	14.13	1.03
advance	391	5,952	97.28	63.08	197	3,035	40.79	23.14	1.03

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
scope	378	5,665	94.04	60.04	181	2,503	37.47	19.09	1.03
presentation	346	4,963	86.08	52.60	146	2,018	30.23	15.39	1.03
behavioral	320	4,377	79.61	46.39	86	1,001	17.81	7.63	1.03
lot	301	3,926	74.89	41.61	128	1,762	26.50	13.44	1.03
suggestion	285	3,589	70.91	38.04	134	2,087	27.74	15.91	1.03
schedule	278	3,407	69.16	36.11	98	1,095	20.29	8.35	1.03
poverty	271	3,237	67.42	34.31	82	715	16.98	5.45	1.03
se	268	3,285	66.68	34.82	131	1,445	27.12	11.02	1.03
clinician	171	1,040	42.54	11.02	21	301	4.35	2.30	1.03
asthma	162	812	40.30	8.61	7	117	1.45	0.89	1.03
facilitation	154	638	38.31	6.76	55	215	11.39	1.64	1.03
danish	151	595	37.57	6.31	16	180	3.31	1.37	1.03
homeless	144	480	35.83	5.09	11	118	2.28	0.90	1.03
dental	142	375	35.33	3.97	6	125	1.24	0.95	1.03
unintended	142	420	35.33	4.45	70	289	14.49	2.20	1.03
to	83,597	1,901,016	20,798.17	20,148.07	482	13,115	99.79	100.00	1.03
get	935	18,439	232.62	195.43	265	5,157	54.87	39.32	1.03
week	554	9,731	137.83	103.13	137	2,810	28.36	21.43	1.03
formal	519	8,959	129.12	94.95	183	2,948	37.89	22.48	1.03
climate	364	5,489	90.56	58.18	86	1,126	17.81	8.59	1.03
cite	345	4,988	85.83	52.87	135	2,461	27.95	18.76	1.03
outline	335	4,810	83.34	50.98	163	2,739	33.75	20.88	1.03
officer	290	3,765	72.15	39.90	64	821	13.25	6.26	1.03
agenda	269	3,289	66.92	34.86	107	1,280	22.15	9.76	1.03
scientist	262	3,111	65.18	32.97	79	993	16.36	7.57	1.03
personnel	208	1,877	51.75	19.89	77	775	15.94	5.91	1.03
consultation	196	1,621	48.76	17.18	82	553	16.98	4.22	1.03
four	1,308	27,063	325.42	286.83	346	8,026	71.64	61.20	1.03
patent	248	2,936	61.70	31.12	17	314	3.52	2.39	1.03
missing	171	1,178	42.54	12.49	62	641	12.84	4.89	1.03
percentile	158	837	39.31	8.87	20	259	4.14	1.97	1.03
oecd	157	857	39.06	9.08	54	240	11.18	1.83	1.03
disadvantaged	149	662	37.07	7.02	41	339	8.49	2.58	1.03
acknowledgement	142	498	35.33	5.28	131	363	27.12	2.77	1.03
counseling	140	451	34.83	4.78	21	160	4.35	1.22	1.03
ces	126	166	31.35	1.76	21	71	4.35	0.54	1.03

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
christie	124	110	30.85	1.17	43	48	8.90	0.37	1.03
evaluand	122	2	30.35	0.02	36	2	7.45	0.02	1.03
female	782	15,126	194.55	160.31	157	2,546	32.51	19.41	1.03
essential	515	9,036	128.13	95.77	231	4,304	47.83	32.82	1.03
medicine	242	2,765	60.21	29.31	96	952	19.88	7.26	1.03
overview	213	2,117	52.99	22.44	129	1,397	26.71	10.65	1.03
weakness	206	1,923	51.25	20.38	105	1,207	21.74	9.20	1.03
consult	169	1,131	42.05	11.99	82	723	16.98	5.51	1.03
shelter	161	875	40.06	9.27	14	402	2.90	3.07	1.03
integrated	251	3,048	62.45	32.30	96	1,367	19.88	10.42	1.03
systematically	210	2,093	52.25	22.18	120	1,405	24.84	10.71	1.03
johnson	192	1,767	47.77	18.73	85	780	17.60	5.95	1.03
workforce	164	1,067	40.80	11.31	56	374	11.59	2.85	1.03
kim	157	920	39.06	9.75	53	362	10.97	2.76	1.03
denmark	153	837	38.06	8.87	35	380	7.25	2.90	1.03
e-mail	153	804	38.06	8.52	114	307	23.60	2.34	1.03
pretest	141	575	35.08	6.09	22	119	4.55	0.91	1.03
overarching	138	445	34.33	4.72	77	342	15.94	2.61	1.03
changes	137	424	34.08	4.49	40	283	8.28	2.16	1.03
metacognitive	130	265	32.34	2.81	11	52	2.28	0.40	1.03
detroit	124	212	30.85	2.25	10	87	2.07	0.66	1.03
funded	122	91	30.35	0.96	38	82	7.87	0.63	1.03
logframe	118	0	29.36	0.00	5	0	1.04	0.00	1.03
why	803	15,704	199.78	166.44	288	5,349	59.63	40.79	1.03
full	725	13,856	180.37	146.85	255	5,483	52.80	41.81	1.03
empirical	644	12,027	160.22	127.47	233	3,198	48.24	24.38	1.03
theoretical	633	11,762	157.48	124.66	217	3,860	44.93	29.43	1.03
aid	327	4,757	81.35	50.42	127	2,022	26.29	15.42	1.03
translate	308	4,383	76.63	46.45	138	2,243	28.57	17.10	1.03
box	292	4,041	72.65	42.83	112	1,496	23.19	11.41	1.03
package	233	2,648	57.97	28.07	77	1,283	15.94	9.78	1.03
predictor	287	3,989	71.40	42.28	71	1,020	14.70	7.78	1.03
follow-up	247	3,036	61.45	32.18	97	877	20.08	6.69	1.03
utilization	177	1,439	44.04	15.25	74	597	15.32	4.55	1.03
cohen	176	1,452	43.79	15.39	69	564	14.29	4.30	1.03
aboriginal	167	1,256	41.55	13.31	15	125	3.11	0.95	1.03
pedagogical	162	1,100	40.30	11.66	49	373	10.14	2.84	1.03
com	162	1,140	40.30	12.08	97	369	20.08	2.81	1.03

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
emergent	158	1,055	39.31	11.18	67	490	13.87	3.74	1.03
empower	157	1,005	39.06	10.65	78	574	16.15	4.38	1.03
hinder	155	992	38.56	10.51	70	759	14.49	5.79	1.03
guiding	136	504	33.84	5.34	75	392	15.53	2.99	1.03
standards	135	498	33.59	5.28	52	250	10.77	1.91	1.03
likert	131	421	32.59	4.46	52	226	10.77	1.72	1.03
doctoral	127	373	31.60	3.95	49	155	10.14	1.18	1.03
teachers	122	258	30.35	2.73	43	135	8.90	1.03	1.03
taker	121	226	30.10	2.40	13	65	2.69	0.50	1.03
mentorship	116	36	28.86	0.38	15	11	3.11	0.08	1.03
alkin	111	2	27.62	0.02	27	1	5.59	0.01	1.03
require	2,010	43,278	500.07	458.69	405	9,663	83.85	73.68	1.03
consideration	545	9,882	135.59	104.74	248	4,313	51.35	32.89	1.03
emphasize	449	7,651	111.71	81.09	198	3,447	40.99	26.28	1.03
translation	414	6,884	103.00	72.96	67	1,534	13.87	11.70	1.03
administration	411	6,804	102.25	72.11	151	2,008	31.26	15.31	1.03
sustain	264	3,484	65.68	36.93	119	1,869	24.64	14.25	1.03
exit	212	2,258	52.74	23.93	30	765	6.21	5.83	1.03
agriculture	195	1,887	48.51	20.00	49	634	10.14	4.83	1.03
lean	152	860	37.82	9.11	18	423	3.73	3.23	1.03
lunch	127	361	31.60	3.83	38	201	7.87	1.53	1.03
english	790	15,555	196.54	164.86	129	2,100	26.71	16.01	1.03
l2	310	4,586	77.13	48.61	22	370	4.55	2.82	1.03
making	193	1,911	48.02	20.25	81	981	16.77	7.48	1.03
developer	171	1,471	42.54	15.59	61	404	12.63	3.08	1.03
im	162	1,201	40.30	12.73	24	260	4.97	1.98	1.03
norwegian	150	910	37.32	9.64	15	182	3.11	1.39	1.03
in-depth	144	847	35.83	8.98	91	565	18.84	4.31	1.03
tailor	144	788	35.83	8.35	94	554	19.46	4.22	1.03
burnout	137	618	34.08	6.55	15	51	3.11	0.39	1.03
preservice	113	139	28.11	1.47	8	19	1.66	0.14	1.03
status	823	16,368	204.75	173.48	218	4,336	45.13	33.06	1.03
original	694	13,413	172.66	142.16	211	4,852	43.69	37.00	1.03
integration	433	7,384	107.73	78.26	102	2,112	21.12	16.10	1.03
office	407	6,800	101.26	72.07	144	2,014	29.81	15.36	1.03
summary	395	6,528	98.27	69.19	183	3,451	37.89	26.31	1.03
motivate	315	4,708	78.37	49.90	147	2,295	30.43	17.50	1.03
equality	289	4,167	71.90	44.16	57	1,156	11.80	8.81	1.03

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
update	272	3,741	67.67	39.65	172	1,135	35.61	8.65	1.03
submit	230	2,782	57.22	29.49	92	1,381	19.05	10.53	1.03
assistant	163	1,244	40.55	13.18	74	626	15.32	4.77	1.03
mail	149	968	37.07	10.26	66	414	13.66	3.16	1.03
sampling	290	4,290	72.15	45.47	111	1,411	22.98	10.76	1.03
mapping	269	3,757	66.92	39.82	64	1,009	13.25	7.69	1.03
co	235	3,024	58.47	32.05	86	498	17.81	3.80	1.03
geographic	195	2,024	48.51	21.45	70	775	14.49	5.91	1.03
self-efficacy	184	1,811	45.78	19.19	34	136	7.04	1.04	1.03
propensity	179	1,661	44.53	17.60	41	633	8.49	4.83	1.03
causality	148	997	36.82	10.57	45	363	9.32	2.77	1.03
challenging	148	1,037	36.82	10.99	107	762	22.15	5.81	1.03
tertiary	136	676	33.84	7.16	32	315	6.63	2.40	1.03
hispanic	135	666	33.59	7.06	43	210	8.90	1.60	1.03
transformational	129	602	32.09	6.38	24	90	4.97	0.69	1.03
lm	125	471	31.10	4.99	9	100	1.86	0.76	1.03
rmsea	119	305	29.61	3.23	24	92	4.97	0.70	1.03
counselor	111	177	27.62	1.88	15	82	3.11	0.63	1.03
category	944	19,186	234.86	203.34	233	4,058	48.24	30.94	1.03
match	562	10,453	139.82	110.79	152	3,130	31.47	23.87	1.03
database	398	6,730	99.02	71.33	101	1,675	20.91	12.77	1.03
vision	314	4,784	78.12	50.70	113	1,677	23.40	12.79	1.03
emotion	305	4,616	75.88	48.92	38	874	7.87	6.66	1.03
consensus	291	4,228	72.40	44.81	151	1,706	31.26	13.01	1.03
psychology	248	3,254	61.70	34.49	80	958	16.56	7.30	1.03
professor	209	2,409	52.00	25.53	57	836	11.80	6.37	1.03
productive	201	2,193	50.01	23.24	67	1,071	13.87	8.17	1.03
vegetable	142	882	35.33	9.35	18	239	3.73	1.82	1.03
interested	300	4,540	74.64	48.12	169	2,675	34.99	20.40	1.03
advanced	228	2,898	56.72	30.71	99	1,578	20.50	12.03	1.03
workplace	201	2,281	50.01	24.18	50	478	10.35	3.64	1.03
certification	128	676	31.85	7.16	37	184	7.66	1.40	1.03
observational	128	642	31.85	6.80	45	287	9.32	2.19	1.03
mu	125	516	31.10	5.47	62	103	12.84	0.79	1.03
remediation	111	231	27.62	2.45	7	72	1.45	0.55	1.03
multidisciplinary	110	217	27.37	2.30	43	141	8.90	1.08	1.03
phd	110	260	27.37	2.76	49	153	10.14	1.17	1.03
efa	107	142	26.62	1.51	14	26	2.90	0.20	1.03

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
donaldson	107	169	26.62	1.79	38	50	7.87	0.38	1.03
theory-based	104	65	25.87	0.69	35	33	7.25	0.25	1.03
dpme	102	0	25.38	0.00	5	0	1.04	0.00	1.03
stufflebeam	101	4	25.13	0.04	21	2	4.35	0.02	1.03
average	1,375	29,202	342.09	309.50	236	5,830	48.86	44.45	1.03
majority	510	9,423	126.88	99.87	208	4,054	43.06	30.91	1.03
commitment	433	7,578	107.73	80.32	183	2,041	37.89	15.56	1.03
australia	264	3,873	65.68	41.05	78	868	16.15	6.62	1.02
mandate	169	1,664	42.05	17.64	77	726	15.94	5.54	1.02
humanity	161	1,512	40.06	16.03	40	635	8.28	4.84	1.02
em	140	976	34.83	10.34	60	231	12.42	1.76	1.02
supplemental	136	917	33.84	9.72	41	278	8.49	2.12	1.02
invariance	133	838	33.09	8.88	11	199	2.28	1.52	1.02
dissatisfaction	125	625	31.10	6.62	14	348	2.90	2.65	1.02
timely	124	677	30.85	7.18	71	447	14.70	3.41	1.02
multilevel	117	502	29.11	5.32	27	146	5.59	1.11	1.02
networking	117	488	29.11	5.17	48	200	9.94	1.52	1.02
semi-structured	111	323	27.62	3.42	51	217	10.56	1.65	1.02
pharmacy	109	302	27.12	3.20	15	81	3.11	0.62	1.02
programmatic	109	314	27.12	3.33	45	161	9.32	1.23	1.02
ou	106	226	26.37	2.40	36	91	7.45	0.69	1.02
rigor	104	196	25.87	2.08	53	145	10.97	1.11	1.02
rasch	101	147	25.13	1.56	9	26	1.86	0.20	1.02
postsecondary	100	52	24.88	0.55	28	24	5.80	0.18	1.02
wic	96	9	23.88	0.10	5	4	1.04	0.03	1.02
result	6,000	135,238	1,492.74	1,433.33	456	12,512	94.41	95.40	1.02
total	1,440	30,802	358.26	326.46	292	7,464	60.46	56.91	1.02
version	688	13,606	171.17	144.20	203	4,580	42.03	34.92	1.02
rural	348	5,759	86.58	61.04	81	1,021	16.77	7.78	1.02
proposal	312	4,971	77.62	52.69	91	1,653	18.84	12.60	1.02
availability	263	3,770	65.43	39.96	115	1,908	23.81	14.55	1.02
advice	201	2,355	50.01	24.96	104	1,020	21.53	7.78	1.02
safe	198	2,351	49.26	24.92	74	1,220	15.32	9.30	1.02
phone	153	1,300	38.06	13.78	74	484	15.32	3.69	1.02
flip	118	495	29.36	5.25	14	234	2.90	1.78	1.02
eight	335	5,523	83.34	58.54	151	2,935	31.26	22.38	1.02
experienced	155	1,482	38.56	15.71	92	827	19.05	6.31	1.02
hs	146	1,208	36.32	12.80	11	140	2.28	1.07	1.02

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
excellence	126	762	31.35	8.08	43	447	8.90	3.41	1.02
texas	121	680	30.10	7.21	21	297	4.35	2.26	1.02
davies	112	486	27.86	5.15	36	218	7.45	1.66	1.02
accreditation	110	368	27.37	3.90	20	93	4.14	0.71	1.02
curricular	105	298	26.12	3.16	31	112	6.42	0.85	1.02
oo	100	161	24.88	1.71	16	21	3.31	0.16	1.02
mum	100	172	24.88	1.82	48	57	9.94	0.43	1.02
ect	95	98	23.64	1.04	34	37	7.04	0.28	1.02
twitter	93	3	23.14	0.03	8	3	1.66	0.02	1.02
family	1,378	29,466	342.83	312.30	196	4,491	40.58	34.24	1.02
target	1,124	23,633	279.64	250.48	258	4,233	53.42	32.28	1.02
month	612	11,870	152.26	125.81	158	3,218	32.71	24.54	1.02
book	593	11,530	147.53	122.20	123	2,537	25.47	19.34	1.02
medical	438	7,942	108.97	84.17	109	1,755	22.57	13.38	1.02
employment	424	7,603	105.49	80.58	114	1,411	23.60	10.76	1.02
positively	294	4,671	73.14	49.51	138	2,018	28.57	15.39	1.02
ownership	273	4,176	67.92	44.26	97	938	20.08	7.15	1.02
coordination	249	3,572	61.95	37.86	69	939	14.29	7.16	1.02
articulate	237	3,328	58.96	35.27	117	1,582	24.22	12.06	1.02
facet	154	1,437	38.31	15.23	35	550	7.25	4.19	1.02
usa	316	5,230	78.62	55.43	118	1,708	24.43	13.02	1.02
beneficial	192	2,382	47.77	25.25	124	1,385	25.67	10.56	1.02
rationale	185	2,233	46.03	23.67	101	1,250	20.91	9.53	1.02
dataset	160	1,651	39.81	17.50	55	469	11.39	3.58	1.02
stratum	143	1,207	35.58	12.79	12	346	2.48	2.64	1.02
underpin	141	1,243	35.08	13.17	64	760	13.25	5.79	1.02
typology	133	993	33.09	10.52	26	380	5.38	2.90	1.02
advocacy	129	972	32.09	10.30	61	321	12.63	2.45	1.02
visualization	119	743	29.61	7.87	28	356	5.80	2.71	1.02
tier	118	704	29.36	7.46	16	185	3.31	1.41	1.02
ee	115	655	28.61	6.94	29	91	6.00	0.69	1.02
sis	110	462	27.37	4.90	29	106	6.00	0.81	1.02
engaging	99	245	24.63	2.60	62	205	12.84	1.56	1.02
tennessee	95	108	23.64	1.14	12	68	2.48	0.52	1.02
health-care	95	129	23.64	1.37	18	43	3.73	0.33	1.02
prioritize	94	161	23.39	1.71	56	119	11.59	0.91	1.02
wellness	92	64	22.89	0.68	16	30	3.31	0.23	1.02
mel	91	68	22.64	0.72	8	34	1.66	0.26	1.02

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Doc Freq	RC Rel Doc Freq	Score
mde	90	1	22.39	0.01	5	1	1.04	0.01	1.02
initial	944	19,662	234.86	208.39	278	6,031	57.56	45.99	1.02
effectively	324	5,377	80.61	56.99	178	3,152	36.85	24.03	1.02
specialist	174	1,985	43.29	21.04	82	873	16.98	6.66	1.02
thank	173	1,963	43.04	20.81	130	1,205	26.92	9.19	1.02
tutor	144	1,268	35.83	13.44	23	201	4.76	1.53	1.02
editor	140	1,161	34.83	12.30	70	588	14.49	4.48	1.02
disaster	133	1,013	33.09	10.74	18	403	3.73	3.07	1.02
pd	157	1,706	39.06	18.08	11	180	2.28	1.37	1.02
moderator	114	650	28.36	6.89	26	170	5.38	1.30	1.02
bloom	103	400	25.63	4.24	25	170	5.18	1.30	1.02
worldview	101	381	25.13	4.04	26	219	5.38	1.67	1.02
constructivist	98	310	24.38	3.29	24	117	4.97	0.89	1.02
ria	90	133	22.39	1.41	7	47	1.45	0.36	1.02
alumnus	90	144	22.39	1.53	9	41	1.86	0.31	1.02
turnaround	88	88	21.89	0.93	10	58	2.07	0.44	1.02
sig	87	92	21.64	0.98	12	27	2.48	0.21	1.02

Appendix C

Evaluation Specific Word List

Item	FC Freq	RC Freq	FC Rel Freq	RC Rel Freq	FC Doc Freq	RC Doc Freq	FC Rel Freq	RC Rel Freq	Score
evaluator	3884	470	966.30	4.98	244	87	50.52	0.66	1.96
et	8565	77138	2130.89	817.55	420	6285	86.96	47.92	1.72
programme	2705	6522	672.98	69.12	122	1453	25.26	11.08	1.57
stakeholder	2406	2194	598.59	23.25	273	474	56.52	3.61	1.56
datum	6642	80212	1652.47	850.13	442	8953	91.51	68.27	1.43
learning	2320	9613	577.19	101.88	261	1471	54.04	11.22	1.43
training	2657	15590	661.04	165.23	287	2242	59.42	17.09	1.43
funding	1837	2477	457.03	26.25	238	818	49.28	6.24	1.42
finding	2919	23557	726.22	249.67	399	5664	82.61	43.19	1.38
planning	1828	7516	454.79	79.66	260	1297	53.83	9.89	1.35
sustainability	1293	1290	321.69	13.67	113	289	23.40	2.20	1.30
teaching	1578	7816	392.59	82.84	160	1266	33.13	9.65	1.29
engagement	1265	3997	314.72	42.36	248	1245	51.35	9.49	1.26
a1	1195	2770	297.31	29.36	201	570	41.61	4.35	1.26
their	12105	211018	3011.61	2236.49	477	12276	98.76	93.60	1.24
collaboration	1047	2438	260.48	25.84	202	877	41.82	6.69	1.23
doi	778	156	193.56	1.65	190	37	39.34	0.28	1.19
rubric	782	533	194.55	5.65	44	237	9.11	1.81	1.19
rater	728	738	181.12	7.82	24	148	4.97	1.13	1.17
competency	729	1244	181.37	13.18	117	341	24.22	2.60	1.17
understanding	1477	16536	367.46	175.26	347	5323	71.84	40.59	1.16
math	688	549	171.17	5.82	50	156	10.35	1.19	1.16
organizational	982	7709	244.31	81.70	155	1016	32.09	7.75	1.15
respondent	1053	9586	261.98	101.60	157	1153	32.51	8.79	1.15
collaborative	639	1764	158.98	18.70	178	525	36.85	4.00	1.14
evaluative	612	1209	152.26	12.81	121	282	25.05	2.15	1.14
causal	776	4928	193.06	52.23	124	1113	25.67	8.49	1.13
monitoring	679	2969	168.93	31.47	129	1054	26.71	8.04	1.13
accountability	621	1770	154.50	18.76	157	481	32.51	3.67	1.13
meeting	791	5387	196.79	57.09	178	1674	36.85	12.76	1.13
participatory	566	748	140.82	7.93	120	243	24.84	1.85	1.13
sd	692	3408	172.16	36.12	113	753	23.40	5.74	1.13
organisation	718	4487	178.63	47.56	96	854	19.88	6.51	1.13
thinking	673	3545	167.44	37.57	175	1431	36.23	10.91	1.13

setting	985	10413	245.06	110.36	276	3481	57.14	26.54	1.12
metric	642	3460	159.72	36.67	82	683	16.98	5.21	1.12
contextual	597	2453	148.53	26.00	163	828	33.75	6.31	1.12
our	4934	94431	1227.53	1000.83	396	10352	81.99	78.93	1.11
literacy	553	2062	137.58	21.85	81	361	16.77	2.75	1.11
toc	455	355	113.20	3.76	17	41	3.52	0.31	1.11
standardized	550	2485	136.83	26.34	126	1080	26.09	8.23	1.11
cjpe	432	0	107.48	0.00	30	0	6.21	0.00	1.11
reading	916	10523	227.89	111.53	101	2303	20.91	17.56	1.11
instructional	481	1300	119.67	13.78	74	271	15.32	2.07	1.10
patton	424	101	105.49	1.07	99	53	20.50	0.40	1.10
implementer	408	35	101.51	0.37	43	29	8.90	0.22	1.10
citation	443	866	110.21	9.18	47	303	9.73	2.31	1.10
charter	474	1592	117.93	16.87	20	300	4.14	2.29	1.10
your	775	8181	192.81	86.71	154	2121	31.88	16.17	1.10
realist	425	776	105.74	8.22	32	195	6.63	1.49	1.10
enrollment	423	667	105.24	7.07	48	175	9.94	1.33	1.10
facilitator	412	479	102.50	5.08	90	190	18.63	1.45	1.10
administrative	540	3312	134.35	35.10	117	1048	24.22	7.99	1.10
align	516	2757	128.38	29.22	208	1367	43.06	10.42	1.10
enroll	438	1086	108.97	11.51	85	490	17.60	3.74	1.10
applicant	441	1433	109.72	15.19	41	230	8.49	1.75	1.09
housing	614	5452	152.76	57.78	43	766	8.90	5.84	1.09
foster	530	3750	131.86	39.74	165	1565	34.16	11.93	1.09
rating	637	6148	158.48	65.16	91	855	18.84	6.52	1.09
provider	518	3688	128.87	39.09	106	720	21.95	5.49	1.09
formative	384	788	95.54	8.35	106	274	21.95	2.09	1.09
ranking	436	2069	108.47	21.93	42	585	8.70	4.46	1.09
learner	589	5486	146.54	58.14	87	450	18.01	3.43	1.08
el	420	1794	104.49	19.01	47	370	9.73	2.82	1.08
educator	410	1638	102.00	17.36	89	419	18.43	3.19	1.08
testing	611	6124	152.01	64.91	145	2069	30.02	15.78	1.08
ongoing	480	3306	119.42	35.04	196	1894	40.58	14.44	1.08
consulting	351	434	87.33	4.60	43	202	8.90	1.54	1.08
governance	527	4503	131.11	47.73	84	686	17.39	5.23	1.08
baseline	524	4427	130.37	46.92	99	1129	20.50	8.61	1.08
canada	460	3006	114.44	31.86	95	886	19.67	6.76	1.08

prevention	409	1955	101.76	20.72	93	709	19.25	5.41	1.08
beneficiary	348	681	86.58	7.22	71	267	14.70	2.04	1.08
shared	384	1621	95.54	17.18	112	785	23.19	5.99	1.08
methodological	413	2324	102.75	24.63	166	1100	34.37	8.39	1.08
institute	448	3230	111.46	34.23	187	1871	38.72	14.27	1.08
demographic	454	3449	112.95	36.55	145	1162	30.02	8.86	1.07
feasibility	333	787	82.85	8.34	68	464	14.08	3.54	1.07
aea	295	46	73.39	0.49	29	10	6.00	0.08	1.07
writing	533	5343	132.61	56.63	86	1631	17.81	12.44	1.07
reflective	354	1438	88.07	15.24	83	606	17.18	4.62	1.07
studies	339	1050	84.34	11.13	115	540	23.81	4.12	1.07
screening	393	2399	97.77	25.43	43	764	8.90	5.83	1.07
ise	293	143	72.90	1.52	8	17	1.66	0.13	1.07
efficacy	417	2970	103.75	31.48	109	1119	22.57	8.53	1.07
cc	314	740	78.12	7.84	20	172	4.14	1.31	1.07
grantee	283	23	70.41	0.24	20	11	4.14	0.08	1.07
decision-making	410	2935	102.00	31.11	136	1003	28.16	7.65	1.07
re	345	1481	85.83	15.70	109	341	22.57	2.60	1.07
pa	371	2168	92.30	22.98	40	522	8.28	3.98	1.07
ev	304	692	75.63	7.33	43	152	8.90	1.16	1.07
descriptive	404	2974	100.51	31.52	160	1399	33.13	10.67	1.07
ea	295	524	73.39	5.55	35	129	7.25	0.98	1.07
fidelity	316	1124	78.62	11.91	60	246	12.42	1.88	1.07
earnings	266	12	66.18	0.13	20	9	4.14	0.07	1.07
limited	621	7968	154.50	84.45	263	4388	54.45	33.46	1.07
instructor	329	1481	81.85	15.70	48	255	9.94	1.94	1.07
ecb	266	234	66.18	2.48	9	20	1.86	0.15	1.06
empowerment	288	808	71.65	8.56	75	271	15.53	2.07	1.06
evidence-based	271	351	67.42	3.72	88	121	18.22	0.92	1.06
remedial	263	215	65.43	2.28	7	121	1.45	0.92	1.06
centre	602	7783	149.77	82.49	119	2213	24.64	16.87	1.06
five	953	15652	237.10	165.89	312	5969	64.60	45.51	1.06
wam	238	22	59.21	0.23	8	3	1.66	0.02	1.06
recruitment	344	2456	85.58	26.03	81	774	16.77	5.90	1.06
guidance	344	2443	85.58	25.89	156	1021	32.30	7.78	1.06
na	295	1344	73.39	14.24	35	437	7.25	3.33	1.06
holistic	269	812	66.92	8.61	73	394	15.11	3.00	1.06

hf	260	623	64.69	6.60	6	133	1.24	1.01	1.06
mindset	243	256	60.46	2.71	32	138	6.63	1.05	1.06
funder	240	133	59.71	1.41	110	53	22.77	0.40	1.06
my	1311	23986	326.16	254.22	191	3345	39.54	25.51	1.06
interdisciplinary	248	401	61.70	4.25	47	215	9.73	1.64	1.06
aej	230	0	57.22	0.00	47	0	9.73	0.00	1.06
me	813	13109	202.27	138.94	220	2749	45.55	20.96	1.06
additionally	395	3789	98.27	40.16	155	2052	32.09	15.65	1.06
innovative	319	2054	79.36	21.77	123	902	25.47	6.88	1.06
sciences	253	642	62.94	6.80	109	421	22.57	3.21	1.06
wow	225	38	55.98	0.40	16	33	3.31	0.25	1.06
covariate	248	569	61.70	6.03	43	175	8.90	1.33	1.06
three	2680	54875	666.76	581.60	424	10746	87.78	81.94	1.05
intended	253	814	62.94	8.63	107	498	22.15	3.80	1.05
dissemination	252	787	62.70	8.34	102	451	21.12	3.44	1.05
completion	319	2405	79.36	25.49	116	1176	24.02	8.97	1.05
coordinator	231	374	57.47	3.96	52	141	10.77	1.08	1.05
rigorous	262	1181	65.18	12.52	121	860	25.05	6.56	1.05
nations	255	996	63.44	10.56	54	403	11.18	3.07	1.05
subscale	242	691	60.21	7.32	32	121	6.63	0.92	1.05
st	344	3118	85.58	33.05	77	717	15.94	5.47	1.05
culturally	276	1585	68.67	16.80	77	753	15.94	5.74	1.05
eligible	246	959	61.20	10.16	72	471	14.91	3.59	1.05
keywords	227	509	56.48	5.39	227	294	47.00	2.24	1.05
readiness	224	420	55.73	4.45	68	223	14.08	1.70	1.05
ministry	314	2568	78.12	27.22	85	644	17.60	4.91	1.05
subgroup	294	2119	73.14	22.46	48	573	9.94	4.37	1.05
organisational	260	1279	64.69	13.56	51	252	10.56	1.92	1.05
responsive	259	1306	64.44	13.84	97	724	20.08	5.52	1.05
supervisor	251	1139	62.45	12.07	47	268	9.73	2.04	1.05
programming	279	1844	69.41	19.54	91	552	18.84	4.21	1.05
counterfactual	234	787	58.22	8.34	39	136	8.07	1.04	1.05
lee	279	1962	69.41	20.79	109	736	22.57	5.61	1.05
website	234	954	58.22	10.11	85	417	17.60	3.18	1.05
nutrition	229	763	56.97	8.09	51	249	10.56	1.90	1.05
attrition	214	490	53.24	5.19	46	185	9.52	1.41	1.05
pre-service	200	138	49.76	1.46	12	28	2.48	0.21	1.05

checklist	210	436	52.25	4.62	60	170	12.42	1.30	1.05
analyse	443	5815	110.21	61.63	114	2328	23.60	17.75	1.05
administer	338	3392	84.09	35.95	141	1518	29.19	11.57	1.05
collaborate	215	644	53.49	6.83	112	431	23.19	3.29	1.05
summative	191	134	47.52	1.42	66	55	13.66	0.42	1.05
canadian	275	2128	68.42	22.55	76	495	15.73	3.77	1.05
reporting	271	2004	67.42	21.24	104	582	21.53	4.44	1.05
validation	260	1767	64.69	18.73	80	739	16.56	5.63	1.05
administrator	239	1288	59.46	13.65	87	521	18.01	3.97	1.05
healthcare	213	766	52.99	8.12	59	272	12.22	2.07	1.05
caregiver	200	440	49.76	4.66	30	123	6.21	0.94	1.05
community-based	199	423	49.51	4.48	65	167	13.46	1.27	1.05
africa	378	4569	94.04	48.42	54	864	11.18	6.59	1.04
relevance	360	4093	89.56	43.38	153	2065	31.68	15.75	1.04
validate	285	2430	70.91	25.75	127	1406	26.29	10.72	1.04
operational	264	2005	65.68	21.25	90	853	18.63	6.50	1.04
reviewer	210	717	52.25	7.60	78	356	16.15	2.71	1.04
low-income	202	572	50.26	6.06	32	176	6.63	1.34	1.04
nursing	198	507	49.26	5.37	33	187	6.83	1.43	1.04
pawson	177	14	44.04	0.15	35	8	7.25	0.06	1.04
existing	405	5233	100.76	55.46	194	2727	40.17	20.79	1.04
aspiration	236	1380	58.71	14.63	45	743	9.32	5.67	1.04
graduation	190	392	47.27	4.15	47	128	9.73	0.98	1.04
efl	185	275	46.03	2.91	9	38	1.86	0.29	1.04
bachelor	184	265	45.78	2.81	48	130	9.94	0.99	1.04
ela	176	60	43.79	0.64	11	9	2.28	0.07	1.04
transdisciplinary	175	51	43.54	0.54	14	15	2.90	0.11	1.04
towards	699	11994	173.90	127.12	183	4209	37.89	32.09	1.04
reasoning	382	4796	95.04	50.83	72	1526	14.91	11.64	1.04
supplementary	243	1650	60.46	17.49	46	527	9.52	4.02	1.04
thematic	213	968	52.99	10.26	85	379	17.60	2.89	1.04
coaching	192	503	47.77	5.33	39	119	8.07	0.91	1.04
mentoring	185	360	46.03	3.82	41	90	8.49	0.69	1.04
bibliometric	170	9	42.29	0.10	24	4	4.97	0.03	1.04
cmo	169	38	42.05	0.40	18	1	3.73	0.01	1.04
dialogue	315	3429	78.37	36.34	87	985	18.01	7.51	1.04
michigan	181	401	45.03	4.25	34	208	7.04	1.59	1.04

nonresponse	171	120	42.54	1.27	15	55	3.11	0.42	1.04
scriven	166	24	41.30	0.25	45	9	9.32	0.07	1.04
building	672	11573	167.19	122.66	199	2263	41.20	17.26	1.04
consistency	313	3431	77.87	36.36	135	1475	27.95	11.25	1.04
suspension	262	2327	65.18	24.66	8	844	1.66	6.44	1.04
investigator	228	1532	56.72	16.24	42	712	8.70	5.43	1.04
refine	227	1531	56.48	16.23	124	1012	25.67	7.72	1.04
ngo	178	400	44.28	4.24	34	136	7.04	1.04	1.04
outreach	171	273	42.54	2.89	43	102	8.90	0.78	1.04
robust	326	3811	81.11	40.39	132	1905	27.33	14.53	1.04
chen	204	1040	50.75	11.02	75	430	15.53	3.28	1.04
cb	172	376	42.79	3.99	7	79	1.45	0.60	1.04
pisa	165	168	41.05	1.78	14	41	2.90	0.31	1.04
professionalization	164	134	40.80	1.42	11	47	2.28	0.36	1.04
well-being	248	2197	61.70	23.29	91	713	18.84	5.44	1.04
analytic	239	1915	59.46	20.30	86	741	17.81	5.65	1.04
systemic	228	1669	56.72	17.69	68	740	14.08	5.64	1.04
proficiency	209	1302	52.00	13.80	53	281	10.97	2.14	1.04
pe	200	1031	49.76	10.93	22	223	4.55	1.70	1.04
keyword	162	164	40.30	1.74	117	68	24.22	0.52	1.04
breastfeed	159	154	39.56	1.63	10	26	2.07	0.20	1.04
cousins	156	35	38.81	0.37	35	19	7.25	0.14	1.04
mayne	153	17	38.06	0.18	24	10	4.97	0.08	1.04
tig	152	9	37.82	0.10	14	3	2.90	0.02	1.04
de	781	14301	194.31	151.57	188	2889	38.92	22.03	1.04
cohort	281	2956	69.91	31.33	54	610	11.18	4.65	1.04
retention	264	2612	65.68	27.68	57	791	11.80	6.03	1.04
infrastructure	264	2646	65.68	28.04	90	794	18.63	6.05	1.04
turnover	250	2284	62.20	24.21	31	619	6.42	4.72	1.04
fairness	207	1335	51.50	14.15	37	364	7.66	2.78	1.04
attainment	204	1243	50.75	13.17	57	510	11.80	3.89	1.04
interviewee	201	1221	50.01	12.94	49	297	10.14	2.26	1.04
targeted	194	1007	48.27	10.67	97	578	20.08	4.41	1.04
retrospective	189	896	47.02	9.50	31	412	6.42	3.14	1.04
eligibility	172	523	42.79	5.54	52	189	10.77	1.44	1.04
psychometric	170	471	42.29	4.99	47	187	9.73	1.43	1.04
homepage	150	33	37.32	0.35	150	23	31.06	0.18	1.04

american	767	14083	190.82	149.26	198	2707	40.99	20.64	1.04
mixed	344	4506	85.58	47.76	137	1996	28.36	15.22	1.04
socioeconomic	196	1152	48.76	12.21	70	467	14.49	3.56	1.04
services	183	912	45.53	9.67	84	481	17.39	3.67	1.04
cronbach	173	643	43.04	6.81	68	270	14.08	2.06	1.04
preschool	169	537	42.05	5.69	22	96	4.55	0.73	1.04
sti	146	54	36.32	0.57	8	17	1.66	0.13	1.04
mentee	143	0	35.58	0.00	7	0	1.45	0.00	1.04
six	611	10703	152.01	113.44	230	4640	47.62	35.38	1.04
supportive	200	1316	49.76	13.95	93	727	19.25	5.54	1.04
campbell	186	1066	46.28	11.30	61	372	12.63	2.84	1.04
exploratory	186	1010	46.28	10.70	81	467	16.77	3.56	1.04
acceptability	166	573	41.30	6.07	24	222	4.97	1.69	1.04
accepted	161	442	40.06	4.68	156	375	32.30	2.86	1.04
condom	159	418	39.56	4.43	8	48	1.66	0.37	1.04
cfa	158	432	39.31	4.58	21	87	4.35	0.66	1.04
um	152	279	37.82	2.96	58	102	12.01	0.78	1.04
pharmacist	145	83	36.07	0.88	10	43	2.07	0.33	1.04
irr	144	92	35.83	0.98	5	13	1.04	0.10	1.04
societal	193	1295	48.02	13.73	62	598	12.84	4.56	1.03
obesity	184	1034	45.78	10.96	26	163	5.38	1.24	1.03
generalizability	159	476	39.56	5.04	54	294	11.18	2.24	1.03
msc	141	93	35.08	0.99	10	28	2.07	0.21	1.03
structured	196	1425	48.76	15.10	91	673	18.84	5.13	1.03
consortium	150	385	37.32	4.08	31	171	6.42	1.30	1.03
maori	149	346	37.07	3.67	8	39	1.66	0.30	1.03
weiss	146	289	36.32	3.06	60	146	12.42	1.11	1.03
app	140	209	34.83	2.22	19	53	3.93	0.40	1.03
heis	138	126	34.33	1.34	7	17	1.45	0.13	1.03
ebp	135	87	33.59	0.92	5	15	1.04	0.11	1.03
united	665	12231	165.45	129.63	240	2560	49.69	19.52	1.03
hiv	220	2086	54.73	22.11	26	202	5.38	1.54	1.03
inspection	209	1796	52.00	19.04	37	973	7.66	7.42	1.03
interpersonal	208	1835	51.75	19.45	70	566	14.49	4.32	1.03
inclusive	190	1372	47.27	14.54	78	652	16.15	4.97	1.03
disciplinary	184	1291	45.78	13.68	53	397	10.97	3.03	1.03
iterative	169	902	42.05	9.56	84	476	17.39	3.63	1.03

policymaker	153	583	38.06	6.18	64	240	13.25	1.83	1.03
rogers	145	391	36.07	4.14	58	203	12.01	1.55	1.03
schools	144	305	35.83	3.23	33	139	6.83	1.06	1.03
ppp	141	297	35.08	3.15	5	47	1.04	0.36	1.03
meta	138	173	34.33	1.83	37	84	7.66	0.64	1.03
rct	137	184	34.08	1.95	30	24	6.21	0.18	1.03
dif	136	187	33.84	1.98	10	33	2.07	0.25	1.03
elsevier	133	107	33.09	1.13	131	39	27.12	0.30	1.03
australasia	131	45	32.59	0.48	29	34	6.00	0.26	1.03
se	268	3285	66.68	34.82	131	1445	27.12	11.02	1.03
clinician	171	1040	42.54	11.02	21	301	4.35	2.30	1.03
asthma	162	812	40.30	8.61	7	117	1.45	0.89	1.03
facilitation	154	638	38.31	6.76	55	215	11.39	1.64	1.03
danish	151	595	37.57	6.31	16	180	3.31	1.37	1.03
homeless	144	480	35.83	5.09	11	118	2.28	0.90	1.03
dental	142	375	35.33	3.97	6	125	1.24	0.95	1.03
unintended	142	420	35.33	4.45	70	289	14.49	2.20	1.03
four	1308	27063	325.42	286.83	346	8026	71.64	61.20	1.03
patent	248	2936	61.70	31.12	17	314	3.52	2.39	1.03
missing	171	1178	42.54	12.49	62	641	12.84	4.89	1.03
percentile	158	837	39.31	8.87	20	259	4.14	1.97	1.03
oecd	157	857	39.06	9.08	54	240	11.18	1.83	1.03
disadvantaged	149	662	37.07	7.02	41	339	8.49	2.58	1.03
acknowledgement	142	498	35.33	5.28	131	363	27.12	2.77	1.03
counseling	140	451	34.83	4.78	21	160	4.35	1.22	1.03
ces	126	166	31.35	1.76	21	71	4.35	0.54	1.03
christie	124	110	30.85	1.17	43	48	8.90	0.37	1.03
evaluand	122	2	30.35	0.02	36	2	7.45	0.02	1.03
integrated	251	3048	62.45	32.30	96	1367	19.88	10.42	1.03
systematically	210	2093	52.25	22.18	120	1405	24.84	10.71	1.03
johnson	192	1767	47.77	18.73	85	780	17.60	5.95	1.03
workforce	164	1067	40.80	11.31	56	374	11.59	2.85	1.03
kim	157	920	39.06	9.75	53	362	10.97	2.76	1.03
denmark	153	837	38.06	8.87	35	380	7.25	2.90	1.03
e-mail	153	804	38.06	8.52	114	307	23.60	2.34	1.03
pretest	141	575	35.08	6.09	22	119	4.55	0.91	1.03
overarching	138	445	34.33	4.72	77	342	15.94	2.61	1.03

changes	137	424	34.08	4.49	40	283	8.28	2.16	1.03
metacognitive	130	265	32.34	2.81	11	52	2.28	0.40	1.03
detroit	124	212	30.85	2.25	10	87	2.07	0.66	1.03
funded	122	91	30.35	0.96	38	82	7.87	0.63	1.03
logframe	118	0	29.36	0.00	5	0	1.04	0.00	1.03
predictor	287	3989	71.40	42.28	71	1020	14.70	7.78	1.03
follow-up	247	3036	61.45	32.18	97	877	20.08	6.69	1.03
utilization	177	1439	44.04	15.25	74	597	15.32	4.55	1.03
cohen	176	1452	43.79	15.39	69	564	14.29	4.30	1.03
aboriginal	167	1256	41.55	13.31	15	125	3.11	0.95	1.03
pedagogical	162	1100	40.30	11.66	49	373	10.14	2.84	1.03
com	162	1140	40.30	12.08	97	369	20.08	2.81	1.03
emergent	158	1055	39.31	11.18	67	490	13.87	3.74	1.03
empower	157	1005	39.06	10.65	78	574	16.15	4.38	1.03
hinder	155	992	38.56	10.51	70	759	14.49	5.79	1.03
guiding	136	504	33.84	5.34	75	392	15.53	2.99	1.03
standards	135	498	33.59	5.28	52	250	10.77	1.91	1.03
likert	131	421	32.59	4.46	52	226	10.77	1.72	1.03
doctoral	127	373	31.60	3.95	49	155	10.14	1.18	1.03
teachers	122	258	30.35	2.73	43	135	8.90	1.03	1.03
taker	121	226	30.10	2.40	13	65	2.69	0.50	1.03
mentorship	116	36	28.86	0.38	15	11	3.11	0.08	1.03
alkin	111	2	27.62	0.02	27	1	5.59	0.01	1.03
english	790	15555	196.54	164.86	129	2100	26.71	16.01	1.03
l2	310	4586	77.13	48.61	22	370	4.55	2.82	1.03
making	193	1911	48.02	20.25	81	981	16.77	7.48	1.03
developer	171	1471	42.54	15.59	61	404	12.63	3.08	1.03
im	162	1201	40.30	12.73	24	260	4.97	1.98	1.03
norwegian	150	910	37.32	9.64	15	182	3.11	1.39	1.03
in-depth	144	847	35.83	8.98	91	565	18.84	4.31	1.03
tailor	144	788	35.83	8.35	94	554	19.46	4.22	1.03
burnout	137	618	34.08	6.55	15	51	3.11	0.39	1.03
preservice	113	139	28.11	1.47	8	19	1.66	0.14	1.03
sampling	290	4290	72.15	45.47	111	1411	22.98	10.76	1.03
mapping	269	3757	66.92	39.82	64	1009	13.25	7.69	1.03
co	235	3024	58.47	32.05	86	498	17.81	3.80	1.03
geographic	195	2024	48.51	21.45	70	775	14.49	5.91	1.03

self-efficacy	184	1811	45.78	19.19	34	136	7.04	1.04	1.03
propensity	179	1661	44.53	17.60	41	633	8.49	4.83	1.03
causality	148	997	36.82	10.57	45	363	9.32	2.77	1.03
challenging	148	1037	36.82	10.99	107	762	22.15	5.81	1.03
tertiary	136	676	33.84	7.16	32	315	6.63	2.40	1.03
hispanic	135	666	33.59	7.06	43	210	8.90	1.60	1.03
transformational	129	602	32.09	6.38	24	90	4.97	0.69	1.03
lm	125	471	31.10	4.99	9	100	1.86	0.76	1.03
rmsea	119	305	29.61	3.23	24	92	4.97	0.70	1.03
counselor	111	177	27.62	1.88	15	82	3.11	0.63	1.03
interested	300	4540	74.64	48.12	169	2675	34.99	20.40	1.03
advanced	228	2898	56.72	30.71	99	1578	20.50	12.03	1.03
workplace	201	2281	50.01	24.18	50	478	10.35	3.64	1.03
certification	128	676	31.85	7.16	37	184	7.66	1.40	1.03
observational	128	642	31.85	6.80	45	287	9.32	2.19	1.03
mu	125	516	31.10	5.47	62	103	12.84	0.79	1.03
remediation	111	231	27.62	2.45	7	72	1.45	0.55	1.03
multidisciplinary	110	217	27.37	2.30	43	141	8.90	1.08	1.03
phd	110	260	27.37	2.76	49	153	10.14	1.17	1.03
efa	107	142	26.62	1.51	14	26	2.90	0.20	1.03
donaldson	107	169	26.62	1.79	38	50	7.87	0.38	1.03
theory-based	104	65	25.87	0.69	35	33	7.25	0.25	1.03
dpme	102	0	25.38	0.00	5	0	1.04	0.00	1.03
stufflebeam	101	4	25.13	0.04	21	2	4.35	0.02	1.03
australia	264	3873	65.68	41.05	78	868	16.15	6.62	1.02
mandate	169	1664	42.05	17.64	77	726	15.94	5.54	1.02
humanity	161	1512	40.06	16.03	40	635	8.28	4.84	1.02
em	140	976	34.83	10.34	60	231	12.42	1.76	1.02
supplemental	136	917	33.84	9.72	41	278	8.49	2.12	1.02
invariance	133	838	33.09	8.88	11	199	2.28	1.52	1.02
dissatisfaction	125	625	31.10	6.62	14	348	2.90	2.65	1.02
timely	124	677	30.85	7.18	71	447	14.70	3.41	1.02
multilevel	117	502	29.11	5.32	27	146	5.59	1.11	1.02
networking	117	488	29.11	5.17	48	200	9.94	1.52	1.02
semi-structured	111	323	27.62	3.42	51	217	10.56	1.65	1.02
pharmacy	109	302	27.12	3.20	15	81	3.11	0.62	1.02
programmatic	109	314	27.12	3.33	45	161	9.32	1.23	1.02

ou	106	226	26.37	2.40	36	91	7.45	0.69	1.02
rigor	104	196	25.87	2.08	53	145	10.97	1.11	1.02
rasch	101	147	25.13	1.56	9	26	1.86	0.20	1.02
postsecondary	100	52	24.88	0.55	28	24	5.80	0.18	1.02
wic	96	9	23.88	0.10	5	4	1.04	0.03	1.02
eight	335	5523	83.34	58.54	151	2935	31.26	22.38	1.02
experienced	155	1482	38.56	15.71	92	827	19.05	6.31	1.02
hs	146	1208	36.32	12.80	11	140	2.28	1.07	1.02
excellence	126	762	31.35	8.08	43	447	8.90	3.41	1.02
texas	121	680	30.10	7.21	21	297	4.35	2.26	1.02
davies	112	486	27.86	5.15	36	218	7.45	1.66	1.02
accreditation	110	368	27.37	3.90	20	93	4.14	0.71	1.02
curricular	105	298	26.12	3.16	31	112	6.42	0.85	1.02
oo	100	161	24.88	1.71	16	21	3.31	0.16	1.02
mum	100	172	24.88	1.82	48	57	9.94	0.43	1.02
ect	95	98	23.64	1.04	34	37	7.04	0.28	1.02
twitter	93	3	23.14	0.03	8	3	1.66	0.02	1.02
usa	316	5230	78.62	55.43	118	1708	24.43	13.02	1.02
beneficial	192	2382	47.77	25.25	124	1385	25.67	10.56	1.02
rationale	185	2233	46.03	23.67	101	1250	20.91	9.53	1.02
dataset	160	1651	39.81	17.50	55	469	11.39	3.58	1.02
stratum	143	1207	35.58	12.79	12	346	2.48	2.64	1.02
underpin	141	1243	35.08	13.17	64	760	13.25	5.79	1.02
typology	133	993	33.09	10.52	26	380	5.38	2.90	1.02
advocacy	129	972	32.09	10.30	61	321	12.63	2.45	1.02
visualization	119	743	29.61	7.87	28	356	5.80	2.71	1.02
tier	118	704	29.36	7.46	16	185	3.31	1.41	1.02
ee	115	655	28.61	6.94	29	91	6.00	0.69	1.02
sis	110	462	27.37	4.90	29	106	6.00	0.81	1.02
engaging	99	245	24.63	2.60	62	205	12.84	1.56	1.02
tennessee	95	108	23.64	1.14	12	68	2.48	0.52	1.02
health-care	95	129	23.64	1.37	18	43	3.73	0.33	1.02
prioritize	94	161	23.39	1.71	56	119	11.59	0.91	1.02
wellness	92	64	22.89	0.68	16	30	3.31	0.23	1.02
mel	91	68	22.64	0.72	8	34	1.66	0.26	1.02
mde	90	1	22.39	0.01	5	1	1.04	0.01	1.02
pd	157	1706	39.06	18.08	11	180	2.28	1.37	1.02

moderator	114	650	28.36	6.89	26	170	5.38	1.30	1.02
bloom	103	400	25.63	4.24	25	170	5.18	1.30	1.02
worldview	101	381	25.13	4.04	26	219	5.38	1.67	1.02
constructivist	98	310	24.38	3.29	24	117	4.97	0.89	1.02
ria	90	133	22.39	1.41	7	47	1.45	0.36	1.02
alumnus	90	144	22.39	1.53	9	41	1.86	0.31	1.02
turnaround	88	88	21.89	0.93	10	58	2.07	0.44	1.02
sig	87	92	21.64	0.98	12	27	2.48	0.21	1.02

Appendix D

Top 500 CEJA2019 Key *n*-grams

Item	Frequency (focus)	Frequency (reference)	Relative frequency (focus)	Relative frequency (reference)	Score
and program planning	908	0	225.90	0.00	1.23
evaluation and program	908	0	225.90	0.00	1.23
evaluation and program planning	904	0	224.91	0.00	1.23
in educational evaluation	653	0	162.46	0.00	1.16
studies in educational	648	0	161.22	0.00	1.16
studies in educational evaluation	648	0	161.22	0.00	1.16
theory of change	643	14	159.97	0.15	1.16
of the program	594	489	147.78	5.18	1.14
of the evaluation	547	185	136.09	1.96	1.13
of the intervention	414	295	103.00	3.13	1.10
new directions for	388	20	96.53	0.21	1.10
directions for evaluation	385	0	95.78	0.00	1.10
new directions for evaluation	385	0	95.78	0.00	1.10
the importance of	727	7678	180.87	81.38	1.09
as well as	1659	28819	412.74	305.44	1.08
in the evaluation	308	259	76.63	2.75	1.07
the implementation of	364	1573	90.56	16.67	1.07
american journal of	273	22	67.92	0.23	1.07
implementation of the	317	1041	78.87	11.03	1.07
journal of evaluation	269	0	66.92	0.00	1.07
american journal of evaluation	269	0	66.92	0.00	1.07
monitoring and evaluation	258	43	64.19	0.46	1.06
the quality of	414	3549	103.00	37.61	1.06
of the research	277	1024	68.92	10.85	1.06
can not be	232	189	57.72	2.00	1.06
as part of	405	4072	100.76	43.16	1.06
of the project	264	929	65.68	9.85	1.06
the purpose of	388	3784	96.53	40.11	1.05
extent to which	376	3667	93.55	38.86	1.05
to participate in	282	1557	70.16	16.50	1.05

the evaluation process	217	53	53.99	0.56	1.05
the use of	937	16263	233.12	172.36	1.05
the impact of	453	5537	112.70	58.68	1.05
the extent to	363	3608	90.31	38.24	1.05
the extent to which	362	3606	90.06	38.22	1.05
the theory of change	195	0	48.51	0.00	1.05
the theory of	248	1339	61.70	14.19	1.05
the health professions	191	6	47.52	0.06	1.05
the evaluation of	227	1051	56.48	11.14	1.05
of the programme	191	189	47.52	2.00	1.05
teaching and learning	203	603	50.50	6.39	1.04
the effectiveness of	259	1907	64.44	20.21	1.04
of the study	293	2753	72.90	29.18	1.04
in the field	264	2167	65.68	22.97	1.04
in the program	179	201	44.53	2.13	1.04
theories of change	169	0	42.05	0.00	1.04
need to be	334	3773	83.10	39.99	1.04
evaluation of the	224	1290	55.73	13.67	1.04
the research team	171	167	42.54	1.77	1.04
in this study	530	8369	131.86	88.70	1.04
understanding of the	344	4122	85.58	43.69	1.04
the field of	245	1907	60.95	20.21	1.04
research and evaluation	160	11	39.81	0.12	1.04
motivation and engagement	159	5	39.56	0.05	1.04
the current study	242	1967	60.21	20.85	1.04
impact of the	212	1375	52.74	14.57	1.04
the control group	205	1162	51.00	12.32	1.04
more likely to	409	5880	101.76	62.32	1.04
it is important	341	4379	84.84	46.41	1.04
the need for	309	3657	76.88	38.76	1.04
included in the	292	3336	72.65	35.36	1.04
to assess the	245	2382	60.95	25.25	1.04
the development of	563	9653	140.07	102.31	1.03
in the study	267	2939	66.43	31.15	1.03
the implementation of the	160	491	39.81	5.20	1.03
to the evaluation	139	85	34.58	0.90	1.03
et al table	138	0	34.33	0.00	1.03

all rights reserved	137	28	34.08	0.30	1.03
the process of	337	4643	83.84	49.21	1.03
is important to	306	3967	76.13	42.04	1.03
a focus on	162	670	40.30	7.10	1.03
lists available at	133	0	33.09	0.00	1.03
contents lists available	133	0	33.09	0.00	1.03
contents lists available at	133	0	33.09	0.00	1.03
a theory of change	132	5	32.84	0.05	1.03
of this study	275	3308	68.42	35.06	1.03
the need to	260	3010	64.69	31.90	1.03
to ensure that	243	2629	60.46	27.86	1.03
the needs of	177	1126	44.04	11.93	1.03
evaluation journal of	130	0	32.34	0.00	1.03
received in revised	129	0	32.09	0.00	1.03
journal of australasia	128	0	31.85	0.00	1.03
in revised form	128	7	31.85	0.07	1.03
received in revised form	128	0	31.85	0.00	1.03
evaluation journal of australasia	128	0	31.85	0.00	1.03
independent evaluation consulting	127	0	31.60	0.00	1.03
based on the	691	12865	171.91	136.35	1.03
validity of the	177	1165	44.04	12.35	1.03
students in the	151	584	37.57	6.19	1.03
on how to	148	540	36.82	5.72	1.03
in higher education	140	332	34.83	3.52	1.03
the program and	129	83	32.09	0.88	1.03
treatment and control	129	136	32.09	1.44	1.03
informal stem education	126	0	31.35	0.00	1.03
the evaluation team	123	6	30.60	0.06	1.03
in terms of	952	18920	236.85	200.53	1.03
in this article	261	3164	64.93	33.53	1.03
a lack of	213	2129	52.99	22.56	1.03
participate in the	168	1086	41.80	11.51	1.03
of an evaluation	121	33	30.10	0.35	1.03
data collection and	129	240	32.09	2.54	1.03

american evaluation association	117	5	29.11	0.05	1.03
in relation to	371	5932	92.30	62.87	1.03
it is important to	273	3666	67.92	38.85	1.03
version of the	234	2739	58.22	29.03	1.03
to improve the	174	1422	43.29	15.07	1.03
a theory of	147	788	36.57	8.35	1.03
participated in the	146	726	36.32	7.69	1.03
to complete the	143	727	35.58	7.71	1.03
participating in the	133	442	33.09	4.68	1.03
the treatment group	119	186	29.61	1.97	1.03
of a program	116	99	28.86	1.05	1.03
professional learning communities	114	11	28.36	0.12	1.03
check for updates	113	0	28.11	0.00	1.03
the logic model	113	0	28.11	0.00	1.03
the findings of	167	1358	41.55	14.39	1.03
purpose of the	160	1146	39.81	12.15	1.03
in the online	121	259	30.10	2.75	1.03
years of experience	118	171	29.36	1.81	1.03
elsevier ltd. all	109	0	27.12	0.00	1.03
ltd. all rights	109	5	27.12	0.05	1.03
elsevier ltd. all rights	109	0	27.12	0.00	1.03
ltd. all rights reserved	109	5	27.12	0.05	1.03
of evaluation in	107	9	26.62	0.10	1.03
in the community	133	666	33.09	7.06	1.03
and implementation of	117	268	29.11	2.84	1.03
knowledge and skills	115	265	28.61	2.81	1.03
we can not	106	42	26.37	0.45	1.03
online version of	105	39	26.12	0.41	1.03
at least one	235	3120	58.47	33.07	1.03
of this article	185	1970	46.03	20.88	1.03
to what extent	148	1127	36.82	11.94	1.03
perceptions of the	133	748	33.09	7.93	1.03
to participate in the	127	631	31.60	6.69	1.03
the online version	104	47	25.87	0.50	1.03
evaluation as a	102	11	25.38	0.12	1.03

in the online version	100	26	24.88	0.28	1.03
a lot of	170	1669	42.29	17.69	1.02
as part of the	169	1638	42.05	17.36	1.02
to focus on	159	1453	39.56	15.40	1.02
the opportunity to	158	1452	39.31	15.39	1.02
participation in the	133	807	33.09	8.55	1.02
of the school	126	670	31.35	7.10	1.02
and how to	113	394	28.11	4.18	1.02
of public health	103	143	25.63	1.52	1.02
the online version of	99	33	24.63	0.35	1.02
field of evaluation	98	0	24.38	0.00	1.02
the field of evaluation	96	0	23.88	0.00	1.02
a range of	265	3935	65.93	41.71	1.02
a total of	226	3047	56.23	32.29	1.02
were asked to	173	1877	43.04	19.89	1.02
would like to	156	1414	38.81	14.99	1.02
based on their	135	988	33.59	10.47	1.02
in the classroom	120	622	29.85	6.59	1.02
of data collection	105	272	26.12	2.88	1.02
at the school	102	215	25.38	2.28	1.02
for the evaluation	102	243	25.38	2.58	1.02
in the outcome	96	106	23.88	1.12	1.02
the evaluation and	95	45	23.64	0.48	1.02
online version of the	93	8	23.14	0.08	1.02
use of evaluation	91	0	22.64	0.00	1.02
the american evaluation	91	0	22.64	0.00	1.02
each of the	387	6873	96.28	72.84	1.02
aspects of the	198	2513	49.26	26.63	1.02
to understand the	187	2250	46.52	23.85	1.02
in the current	185	2198	46.03	23.30	1.02
the purpose of the	122	817	30.35	8.66	1.02
in the research	108	415	26.87	4.40	1.02
in line with	186	2343	46.28	24.83	1.02
this study was	153	1601	38.06	16.97	1.02
the benefits of	145	1347	36.07	14.28	1.02
to develop a	141	1276	35.08	13.52	1.02
the impact of the	128	990	31.85	10.49	1.02

the program is	90	95	22.39	1.01	1.02
to the program	90	95	22.39	1.01	1.02
of an intervention	87	70	21.64	0.74	1.02
evaluation in the	86	30	21.40	0.32	1.02
the intervention and	85	41	21.15	0.43	1.02
the majority of	260	4153	64.69	44.02	1.02
were able to	178	2219	44.28	23.52	1.02
the success of	143	1485	35.58	15.74	1.02
to address the	131	1169	32.59	12.39	1.02
in the field of	126	1006	31.35	10.66	1.02
needs of the	103	560	25.63	5.94	1.02
of the training	101	437	25.13	4.63	1.02
an evaluation of	93	264	23.14	2.80	1.02
and control groups	88	214	21.89	2.27	1.02
the evaluation center	82	0	20.40	0.00	1.02
to the intervention	82	80	20.40	0.85	1.02
independent evaluation consultants	81	0	20.15	0.00	1.02
of evaluation use	80	0	19.90	0.00	1.02
the american evaluation association	80	0	19.90	0.00	1.02
of australasia i	79	0	19.65	0.00	1.02
journal of australasia i	79	0	19.65	0.00	1.02
focused on the	152	1742	37.82	18.46	1.02
the inclusion of	134	1362	33.34	14.44	1.02
to contribute to	116	885	28.86	9.38	1.02
the development and	101	521	25.13	5.52	1.02
of the survey	100	500	24.88	5.30	1.02
the outcomes of	94	446	23.39	4.73	1.02
of the treatment	89	282	22.14	2.99	1.02
quantitative and qualitative	83	179	20.65	1.90	1.02
the program was	81	93	20.15	0.99	1.02
of the journal	80	112	19.90	1.19	1.02
ethnic minority groups	77	48	19.16	0.51	1.02
et al evaluation	75	0	18.66	0.00	1.02
the context of	385	7226	95.78	76.59	1.02
were used to	183	2574	45.53	27.28	1.02

in the process	166	2145	41.30	22.73	1.02
the university of	159	1955	39.56	20.72	1.02
to evaluate the	158	2017	39.31	21.38	1.02
quality of the	147	1726	36.57	18.29	1.02
degree to which	138	1481	34.33	15.70	1.02
the degree to	136	1426	33.84	15.11	1.02
the degree to which	136	1424	33.84	15.09	1.02
shown in table	131	1365	32.59	14.47	1.02
to support the	127	1223	31.60	12.96	1.02
to respond to	120	1120	29.85	11.87	1.02
for this study	113	953	28.11	10.10	1.02
what are the	105	749	26.12	7.94	1.02
and how they	99	627	24.63	6.65	1.02
data were collected	93	519	23.14	5.50	1.02
in the school	91	459	22.64	4.86	1.02
reliability of the	91	476	22.64	5.04	1.02
of the findings	89	412	22.14	4.37	1.02
in the project	86	300	21.40	3.18	1.02
how and why	84	237	20.90	2.51	1.02
evaluation of a	81	160	20.15	1.70	1.02
qualitative and quantitative	80	159	19.90	1.69	1.02
in the intervention	79	107	19.65	1.13	1.02
article reuse guidelines	74	0	18.41	0.00	1.02
evaluation capacity building	73	0	18.16	0.00	1.02
in math and	73	15	18.16	0.16	1.02
of evaluation and	73	42	18.16	0.45	1.02
the intervention was	73	57	18.16	0.60	1.02
teacher evaluation system	72	0	17.91	0.00	1.02
al evaluation and	72	0	17.91	0.00	1.02
et al evaluation and	72	0	17.91	0.00	1.02
al evaluation and program	72	0	17.91	0.00	1.02
program planning journal	71	0	17.66	0.00	1.02
planning journal homepage	71	0	17.66	0.00	1.02
program planning journal homepage	71	0	17.66	0.00	1.02
and program planning journal	71	0	17.66	0.00	1.02

of the data	175	2449	43.54	25.96	1.02
development of the	172	2363	42.79	25.04	1.02
assessment of the	117	1088	29.11	11.53	1.02
of the process	117	1110	29.11	11.76	1.02
the types of	115	1038	28.61	11.00	1.02
review of the	112	976	27.86	10.34	1.02
an opportunity to	108	896	26.87	9.50	1.02
and use of	105	851	26.12	9.02	1.02
the assessment of	100	708	24.88	7.50	1.02
the case study	91	484	22.64	5.13	1.02
to work with	91	526	22.64	5.57	1.02
to better understand	90	485	22.39	5.14	1.02
a case study	85	354	21.15	3.75	1.02
about how to	84	402	20.90	4.26	1.02
to reflect on	80	277	19.90	2.94	1.02
interviews were conducted	79	268	19.65	2.84	1.02
outcomes of the	78	194	19.41	2.06	1.02
who participated in	78	253	19.41	2.68	1.02
department of education	76	168	18.91	1.78	1.02
collection and analysis	75	132	18.66	1.40	1.02
measurement error in	73	82	18.16	0.87	1.02
in an evaluation	71	29	17.66	0.31	1.02
the implementation process	71	33	17.66	0.35	1.02
math and science	71	36	17.66	0.38	1.02
classroom disorder oss	70	0	17.42	0.00	1.02
reflection and feedback	70	0	17.42	0.00	1.02
professional identity tensions	69	0	17.17	0.00	1.02
version of the journal	69	0	17.17	0.00	1.02
from ethnic minority	67	13	16.67	0.14	1.02
well as the	322	5993	80.11	63.52	1.02
as well as the	322	5959	80.11	63.16	1.02
may not be	213	3422	52.99	36.27	1.02
focus on the	184	2797	45.78	29.64	1.02
different types of	153	2094	38.06	22.19	1.02
less likely to	138	1725	34.33	18.28	1.02
the validity of	134	1601	33.34	16.97	1.02

the potential to	118	1239	29.36	13.13	1.02
to explore the	117	1260	29.11	13.35	1.02
a need for	91	638	22.64	6.76	1.02
to develop the	82	435	20.40	4.61	1.02
to implement the	80	383	19.90	4.06	1.02
the design and	80	383	19.90	4.06	1.02
of the course	79	350	19.65	3.71	1.02
of higher education	75	292	18.66	3.09	1.02
how to use	74	250	18.41	2.65	1.02
ministry of education	73	201	18.16	2.13	1.02
of change and	72	151	17.91	1.60	1.02
of research and	72	158	17.91	1.67	1.02
funded by the	72	160	17.91	1.70	1.02
number of students	71	185	17.66	1.96	1.02
the treatment effect	69	89	17.17	0.94	1.02
design and implementation	69	111	17.17	1.18	1.02
the research process	69	142	17.17	1.51	1.02
data collection and analysis	69	93	17.17	0.99	1.02
that the evaluation	68	41	16.92	0.43	1.02
students who are	68	129	16.92	1.37	1.02
from ethnic minority groups	66	5	16.42	0.05	1.02
of evaluative metrics	65	0	16.17	0.00	1.02
the programme theory	64	16	15.92	0.17	1.02
in informal stem	63	0	15.67	0.00	1.02
the results of	383	7438	95.29	78.83	1.02
related to the	335	6382	83.34	67.64	1.02
involved in the	216	3622	53.74	38.39	1.02
characteristics of the	162	2324	40.30	24.63	1.02
to identify the	142	1875	35.33	19.87	1.02
the design of	132	1630	32.84	17.28	1.02
provided by the	128	1582	31.85	16.77	1.02
data from the	123	1419	30.60	15.04	1.02
this study is	111	1162	27.62	12.32	1.02
can lead to	111	1205	27.62	12.77	1.02
on their own	100	918	24.88	9.73	1.02
a review of	93	802	23.14	8.50	1.02
of the impact	84	513	20.90	5.44	1.02

internal and external	80	494	19.90	5.24	1.02
of health and	74	304	18.41	3.22	1.02
the findings from	74	304	18.41	3.22	1.02
of students in	71	212	17.66	2.25	1.02
the impacts of	71	272	17.66	2.88	1.02
strengths and weaknesses	70	240	17.42	2.54	1.02
schools in the	69	236	17.17	2.50	1.02
findings from the	69	240	17.17	2.54	1.02
in the implementation	67	200	16.67	2.12	1.02
in this issue	64	105	15.92	1.11	1.02
like to thank	64	132	15.92	1.40	1.02
on student achievement	63	26	15.67	0.28	1.02
in high school	63	87	15.67	0.92	1.02
measurement error in the	62	36	15.43	0.38	1.02
students'' perceptions of	60	0	14.93	0.00	1.02
corresponding author at	60	0	14.93	0.00	1.02
of principal turnover	60	0	14.93	0.00	1.02
in informal stem education	60	0	14.93	0.00	1.02
culture of impact	59	0	14.68	0.00	1.02
research and innovation	59	13	14.68	0.14	1.02
be able to	262	4773	65.18	50.59	1.01
the lack of	252	4492	62.70	47.61	1.01
in the literature	210	3520	52.25	37.31	1.01
needs to be	178	2838	44.28	30.08	1.01
are more likely to	152	2244	37.82	23.78	1.01
one or more	134	1753	33.34	18.58	1.01
to engage in	124	1569	30.85	16.63	1.01
the characteristics of	122	1498	30.35	15.88	1.01
they did not	121	1538	30.10	16.30	1.01
i do n''t	116	1394	28.86	14.77	1.01
a sample of	110	1212	27.37	12.85	1.01
the quality of the	108	1188	26.87	12.59	1.01
participants in the	106	1118	26.37	11.85	1.01
this article is	100	1054	24.88	11.17	1.01
the department of	90	807	22.39	8.55	1.01
the study was	87	735	21.64	7.79	1.01
the ministry of	84	679	20.90	7.20	1.01

a need to	79	535	19.65	5.67	1.01
of the items	75	402	18.66	4.26	1.01
of the policy	74	366	18.41	3.88	1.01
findings of the	72	410	17.91	4.35	1.01
the needs of the	71	348	17.66	3.69	1.01
theory and practice	69	256	17.17	2.71	1.01
study was conducted	69	300	17.17	3.18	1.01
error in the	68	268	16.92	2.84	1.01
reliability and validity	67	214	16.67	2.27	1.01
the first author	66	215	16.42	2.28	1.01
skills and knowledge	64	133	15.92	1.41	1.01
the focus group	63	129	15.67	1.37	1.01
the data collection	63	200	15.67	2.12	1.01
of the projects	62	106	15.43	1.12	1.01
and professional development	60	59	14.93	0.63	1.01
the sustainability of	60	109	14.93	1.16	1.01
the research and	60	125	14.93	1.32	1.01
critical thinking skills	59	30	14.68	0.32	1.01
on the evaluation	59	72	14.68	0.76	1.01
of program evaluation	58	0	14.43	0.00	1.01
evaluation in informal	58	0	14.43	0.00	1.01
treatment and control groups	58	53	14.43	0.56	1.01
group concept mapping	57	0	14.18	0.00	1.01
evaluation theory and	57	0	14.18	0.00	1.01
approaches to evaluation	57	6	14.18	0.06	1.01
and engagement in	57	32	14.18	0.34	1.01
al studies in	56	0	13.93	0.00	1.01
et al studies	56	0	13.93	0.00	1.01
et al studies in	56	0	13.93	0.00	1.01
al studies in educational	56	0	13.93	0.00	1.01
for evaluators to	55	0	13.68	0.00	1.01
the action plans	55	0	13.68	0.00	1.01
a logic model	55	6	13.68	0.06	1.01
evaluation in informal stem	55	0	13.68	0.00	1.01
was used to	232	4196	57.72	44.47	1.01

the type of	167	2696	41.55	28.57	1.01
are more likely	162	2501	40.30	26.51	1.01
the complexity of	125	1644	31.10	17.42	1.01
the focus of	125	1705	31.10	18.07	1.01
that could be	119	1522	29.61	16.13	1.01
purpose of this	108	1320	26.87	13.99	1.01
the literature on	103	1154	25.63	12.23	1.01
what is the	102	1159	25.38	12.28	1.01
focuses on the	101	1099	25.13	11.65	1.01
to meet the	94	934	23.39	9.90	1.01
in the current study	89	866	22.14	9.18	1.01
an overview of	84	726	20.90	7.69	1.01
do n''t know	83	666	20.65	7.06	1.01
effectiveness of the	81	658	20.15	6.97	1.01
of the literature	79	645	19.65	6.84	1.01
the achievement of	72	440	17.91	4.66	1.01
of change in	69	383	17.17	4.06	1.01
the delivery of	69	386	17.17	4.09	1.01
and how it	69	402	17.17	4.26	1.01
was conducted in	68	344	16.92	3.65	1.01
to identify and	68	390	16.92	4.13	1.01
of the authors	63	293	15.67	3.11	1.01
aligned with the	62	236	15.43	2.50	1.01
confirmatory factor analysis	62	260	15.43	2.76	1.01
to engage with	62	260	15.43	2.76	1.01
and evaluation of	61	184	15.18	1.95	1.01
of the instrument	61	203	15.18	2.15	1.01
health and social	60	153	14.93	1.62	1.01
of the assessment	60	176	14.93	1.87	1.01
of the implementation	60	187	14.93	1.98	1.01
to guide the	60	211	14.93	2.24	1.01
to the research	59	186	14.68	1.97	1.01
department of health	59	187	14.68	1.98	1.01
a culture of	59	201	14.68	2.13	1.01
the school level	57	87	14.18	0.92	1.01
the treatment and	57	111	14.18	1.18	1.01
for the program	56	59	13.93	0.63	1.01

in the programme	56	65	13.93	0.69	1.01
stakeholders in the	56	79	13.93	0.84	1.01
by the program	56	82	13.93	0.87	1.01
and outcomes of	56	87	13.93	0.92	1.01
the school year	56	132	13.93	1.40	1.01
evaluation is a	55	22	13.68	0.23	1.01
that the program	55	90	13.68	0.95	1.01
of the logic	55	92	13.68	0.98	1.01
the data team	54	0	13.43	0.00	1.01
the ise field	54	0	13.43	0.00	1.01
i i i	54	31	13.43	0.33	1.01
the program to	54	68	13.43	0.72	1.01
development and implementation	54	72	13.43	0.76	1.01
standardized achievement tests	53	0	13.19	0.00	1.01
of change for	53	23	13.19	0.24	1.01
error in the outcome	53	0	13.19	0.00	1.01
belcher et al.	52	0	12.94	0.00	1.01
of evaluation as	52	4	12.94	0.04	1.01
health and wellness	52	19	12.94	0.20	1.01
of evaluation systems	51	0	12.69	0.00	1.01
the evaluation field	51	0	12.69	0.00	1.01
of evaluation is	51	7	12.69	0.07	1.01
the aim of	128	1853	31.85	19.64	1.01
the development of the	109	1398	27.12	14.82	1.01
the purpose of this	97	1155	24.13	12.24	1.01
ensure that the	95	1039	23.64	11.01	1.01
who did not	93	1030	23.14	10.92	1.01
a way to	93	1032	23.14	10.94	1.01
presented in table	85	884	21.15	9.37	1.01
to ensure that the	84	818	20.90	8.67	1.01
over the course	83	790	20.65	8.37	1.01
extent to which the	83	770	20.65	8.16	1.01
over the course of	83	787	20.65	8.34	1.01
of the students	77	684	19.16	7.25	1.01
focus of the	74	597	18.41	6.33	1.01

to understand how	74	604	18.41	6.40	1.01
we were able	73	541	18.16	5.73	1.01
of the respondents	73	547	18.16	5.80	1.01
we were able to	73	536	18.16	5.68	1.01
in the treatment	72	585	17.91	6.20	1.01
is a need	69	478	17.17	5.07	1.01
there is a need	69	464	17.17	4.92	1.01
the social sciences	68	412	16.92	4.37	1.01
at the individual	65	401	16.17	4.25	1.01
randomly assigned to	63	389	15.67	4.12	1.01
developed by the	61	303	15.18	3.21	1.01
to develop and	60	302	14.93	3.20	1.01
teachers in the	58	196	14.43	2.08	1.01
the challenges of	58	231	14.43	2.45	1.01
positive effects on	56	134	13.93	1.42	1.01
with a focus	56	170	13.93	1.80	1.01
of the curriculum	56	179	13.93	1.90	1.01
of the framework	55	178	13.68	1.89	1.01
validity and reliability	54	108	13.43	1.14	1.01
would like to thank	54	108	13.43	1.14	1.01
social sciences and	53	71	13.19	0.75	1.01
the project team	53	110	13.19	1.17	1.01
findings from this	53	141	13.19	1.49	1.01
policy and practice	53	146	13.19	1.55	1.01
did not meet	52	134	12.94	1.42	1.01
with mental illness	51	31	12.69	0.33	1.01
after the intervention	51	59	12.69	0.63	1.01
roles and responsibilities	51	78	12.69	0.83	1.01
available at studies	50	0	12.44	0.00	1.01
at studies in	50	0	12.44	0.00	1.01
of assessment quality	50	0	12.44	0.00	1.01
types of evaluation	50	6	12.44	0.06	1.01
and the evaluation	50	52	12.44	0.55	1.01
public health and	50	70	12.44	0.74	1.01
available at studies in	50	0	12.44	0.00	1.01
lists available at studies	50	0	12.44	0.00	1.01
at studies in educational	50	0	12.44	0.00	1.01

approach to evaluation	49	0	12.19	0.00	1.01
the program theory	49	4	12.19	0.04	1.01
optimal sample allocation	48	0	11.94	0.00	1.01