

Trends in Gambling Studies Research: Quantifying, Categorizing, and Describing Citations

Howard J. Shaffer · Michael V. Stanton ·
Sarah E. Nelson

Published online: 16 August 2006
© Springer Science+Business Media, Inc. 2006

Abstract As opportunities to gamble have increased during the 20th century, so has gambling research. This study used new strategies, methods, and technology to examine citation trends and the growth of knowledge in the field of gambling studies. The sample included 2,246 citations that were published between 1903 and 2003. By using multiple keywords to classify each citation into distinct topic areas, this study yielded a more comprehensive analysis than was previously available. The results reveal that gambling-related research has grown at an exponential rate. The most prevalent topics explored within gambling studies citations have been pathology, risk-taking, decision-making and addiction. Between 1999 and 2003, studies addressing epidemiology, drug abuse, comorbidity and neuroscience have become increasingly prevalent. Based on these trends and their implications, this paper provides several recommendations for both future areas of inquiry within the field of gambling studies and better classification techniques for citations within all fields of psychology.

Keywords Gambling · Scientific communication · Technology · Trends · Review

Introduction

Throughout recorded history, people have played many different games of chance; typically, access to these games was restricted by geography and legal status. During the 20th century, however, there was a remarkable growth in the visibility, availability and accessibility of gambling through the proliferation of legalized gambling

H. J. Shaffer (✉) · M. V. Stanton · S. E. Nelson
Division on Addictions, Harvard Medical School, 101 Station Landing, Medford,
MA 02155, USA
e-mail: howard_shaffer@hms.harvard.edu

S. E. Nelson
e-mail: sarah_nelson@hms.harvard.edu

(National Gambling Impact Study Commission, 1999; National Research Council, 1999). Although researchers have written much about gambling and those with gambling-related problems, with one exception (Eber & Shaffer, 2000), researchers have not empirically explored the citation trends and the growth of knowledge in the field of gambling studies. Media attention, the number of related journals, and the number of researchers focusing on gambling studies reflect the size of the field, but these indices do not quantify the extent of gambling-related research and the corresponding growth in gambling-related knowledge.

General citation-based research has been fruitful and provides a perspective that other assessments of the literature cannot provide. For example, Adair and Vohra (2003) note the recent trend of increased references within published psychology articles. They suggest that the use of numerous references has become an important measure of credibility for authors. The growing importance of citations in scholarly publications also reflects an explosion in the number of scientific journal articles that are readily available for citation.

In broad areas of inquiry more mature than gambling studies (e.g., psychology, biology, political science, philosophy, sociology, etc.), the gross number of published scientific and scholarly studies are increasing rapidly (Najman & Hewitt, 2003). The number of psychological publications, specifically, has grown enormously throughout the 20th century (Henderson, 1998; Thorngate, 1990; Xhingnesse & Osgood, 1967). Observation suggests that many fields, such as gambling studies, which draw from these more mature and broadly based disciplines, might show parallel increases in the number of published articles. Quantitative information about the progress of a field can serve as an indicator of the academic attention and resources (e.g., funding) that are dedicated to certain topics of interest.

Evidence about the expansion of an academic field does not necessarily indicate the scholarly consumption of higher quality information. The greater the volume of scholarly literature, the more difficult it becomes to distinguish and access important information (Garfield, 1955; Thorngate, 1990). Classifying information about a field of study helps investigators to identify and prioritize the areas of scientific interest within that field. By organizing citations, researchers can observe empirically areas of investigation that attract more or less attention. By understanding the shifting landscape of scientific research and how scientists describe the constructs under investigation (e.g., nomenclature, theoretical constructs, etc.), it becomes possible to identify and recommend areas of inquiry that deserve more or less attention in the future.

Eber and Shaffer designed the first study of gambling research trends “to quantify psychosocial and biomedical research patterns in gambling studies...” (Eber & Shaffer, 2000, p. 462). They documented the progress of gambling-related studies and also categorized gambling-related publications into subject areas. This research revealed that the field of gambling studies experienced dramatic growth between 1964 and 1999 (i.e., from only a few publications in 1964 to more than 60 publications a year during the late 1990s); in addition, these published studies focused most commonly on the topics “Cognition or Personality” and “Methods or Theory.”

This first empirical analysis of scholarly gambling-related publications relied on (a) heuristic speculation (i.e., using personal judgment to classify citations and create category names), (b) citation categorization based solely on the “title” field, and (c) the restricted categorization of citations by subject because the investigators placed each citation in only one subject category. As a result, Eber and Shaffer (2000) obtained data that relied on an individual coder, restricted information about each

article, and simplified citation subject matter. Because Eber and Shaffer did not extract sufficient topic information about each citation or allow a citation to be grouped into more than one category, their effort to catalog gambling research was unable to classify 21% of the citations. While Eber and Shaffer's restricted analysis provided a concise, unidimensional picture of citation trends, complex citations (i.e., those which explore more than one topic) were necessarily simplified (e.g., an article about epidemiological assessment was categorized as *either* epidemiology *or* methods).

New software (e.g., OVID 9.0, Endnote 7.0, RefViz 1.0, SPSS 12.0) that was previously unavailable, now permits a seamless conversion of citation evidence between programs and more flexible data management. For example, citation data is now cleaned using OVID "deduping" procedures described in more detail below. Moreover, now investigators can translate data more easily between OVID, Endnote, and SPSS, thereby reducing the potential for data management problems. We can now categorize citations more objectively and comprehensively using the following methods: (a) systematic keyword categorization, (b) primary subject-based analysis, and (c) potential classification of citations into multiple topic categories. The goals of this research are to increase the objectivity of citation classification, provide a replicable classification process, and categorize citations more comprehensively. The "keyword" field is used as a benchmark to compare themes in references because it: (a) represents an author's own description of his/her article, (b) contains salient subject terms, and (c) presents topics in a concise, orderly manner. For these reasons, keywords provide an ideal resource for categorizing citations by subject area. In addition, more current data from publications released between 1999 and 2003 will provide the opportunity to confirm if gambling studies have indeed continued to grow in the manner that Eber and Shaffer (2000) observed previously.

Method

Sample

The total sample included 2,246 citations (i.e., scholarly journal articles). To be included, the citations had to satisfy three inclusion criteria: (a) represent articles published between 1903 and 2003, (b) have the word "gambling" in one of three citation fields—the "title" field, the "keyword" field or the "abstract" field, and (c) have some relevance to the field of gambling studies. We excluded citations from the sample that had the word "gambling" in the title or abstract when these articles were unrelated to gambling (e.g., "HIV-1: Gambling on the evolution of drug resistance?" or "The AMA: Gambling with nursing's future" (Brown & Richman, 1997; Rait, 1989)).

Procedure

We conducted literature searches using MEDLINE and PsycINFO databases. A product of the U.S. National Library of Medicine, MEDLINE contains over 11,800,000 references taken from medical journals around the world (U.S. National Library of Medicine, 2004). About 520,000 new references are added annually to the database, which spans from 1966 to the present. A product of the American Psychological Association, PsycINFO contains over 2,000,000 references taken from

psychological journals around the world (American Psychological Association, 2004). About 80,000 new and old references are added annually to the database, which spans from 1872 to the present. We searched both databases through the OVID interface over the Internet (OVID Technologies, 2003).

The present analyses utilize data collected on January 1, 2004. To generate this data, we searched for the term “gambling” in the “title,” “keyword,” or “abstract” citation field of journal articles in both databases simultaneously and removed duplicate citations between the databases through a process called “deduping.” When duplicates were found, we prioritized citations with keywords over those missing this information because keywords were the units of analysis for the categorization procedure; when each version of a duplicate contained different keywords, PsycINFO citations were prioritized over MEDLINE citations. Using the Direct Export feature in OVID and a PsycInfo (OVID) importing filter, we exported these citations into an Endnote database used for managing and collecting citations (Thomson ISI ResearchSoft, 2003).

OVID databases occasionally have more than one entry for the same publication. Therefore, to further remove duplications that were not flagged by the OVID deduping interface, we used the Endnote “find duplicates” function, set to use only “title” and “year” fields as discriminating factors. References with “gambling” in the keyword field were retained but articles with “gambling” in only the title or the abstract fields were reviewed in their entirety for relevancy to gambling-related research. As noted before, we excluded citations referring to the word “gambling” as a metaphor unrelated to gambling studies from further analysis. This process could only be undertaken manually because this task required a method for detecting diminutive differences in citation content.

To identify the growth trends in this data set, we analyzed the frequency of gambling citations across time. To measure content and identify reliable findings that co-occur across differing methods, we employed two different analytic techniques. All three analyses are specified below.

Frequency Analysis Procedure

We exported the data into the *Statistical Package for the Social Sciences* (SPSS Inc., 2003) for frequency analysis. To investigate the change in prevalence of gambling citations over time, we used the citation as the unit of analysis. Data from the year 2003, which might not have been fully archived by our January 1st, 2004 retrieval date, were included in a regression analysis for frequency of citations between 1903 and 2003. However, 2003 data were excluded from the regression for frequency of citations between 1999 and 2002 because incomplete 2003 citations might misrepresent the pattern of growth within a small (i.e., 5-year) sample.

Categorical Analysis Procedure

To analyze category types, we employed the following procedure: the Endnote citation library was exported via .xml format into Microsoft Excel (Microsoft Corporation, 2002) and then from Excel to SPSS. SPSS’s “restructure” procedure was used to organize the file by citation. PsycINFO and MEDLINE database managers (i.e., American Psychological Association and the National Library of Medicine) designate keywords as *primary* keywords by inclusion of an asterisk (G. Hurley,

personal communication, February 27th, 2004). Keywords that are *not* asterisked tend to be more common and generic (e.g., “human”) than these primary keywords. To reduce our data to key topics, we deleted non-asterisked keywords within citations that contained at least one asterisked keyword. For those citations that did not contain *any* asterisked keywords (i.e., those that were published before the primary keyword system was established), we treated all existing keywords as primary keywords. As described below, we used this composite database of primary keywords for two different categorical analyses of citation subject matter. Three citations from the early 1900s contained no keywords; consequently, we did not categorize these citations in either analysis.

Categorical Analysis by Keyword

For the first categorical analysis, we used the composite keyword database to generate frequency counts and create an Excel file with keywords, instead of citations, as the unit of analysis. This new file contained 1,663 different primary keywords. Attached to primary keywords (e.g., gambling, aging, suicide), the composite database originally included 36 different two-letter abbreviations and secondary subject matter (e.g., “/px [psychology],” “/mt [methods],” “/ph [physiology]”). By collapsing across this secondary subject matter (e.g., “addictive behavior/px” and “addictive behavior/pp” became the same keyword), we were able to reduce the number of primary keywords for analysis to 1,442. In this analysis, we used keyword as the unit of analysis and number of citations containing that keyword as the dependent variable. To reduce the primary keywords to a manageable number, keywords that appeared in fewer than 2% of citations were combined into an “other” category. After completing these procedures with the entire sample, we examined recent trends by repeating this analysis of keywords for only those articles published between 1999 and 2003.

Categorical Analysis Using RefViz

For the second categorical analysis, we used RefViz, a citation analysis program that clusters citations according to shared linguistic context using a mathematical technique similar to cluster analysis (Thomson ISI ResearchSoft, 2004), to create groups based on primary keyword terms. We exported the composite database of primary keywords (i.e., keywords preceded by an asterisk for citations that contained at least one asterisked keyword and all keywords from citations without an asterisked keyword) into a tab delimited .txt file and imported this .txt file into Endnote. From Endnote, we exported the file to RefViz using the RefMan (RIS) export method. RefViz conducts its analysis at the level of keyword *terms* (ignoring uninformative articles of speech such as “a,” “the,” and “of” and conflating terms with standard suffixes, such as “gamble” and “gambling”), not keyword *phrases*. For example, RefViz analyzes the keyword “addictive behavior” as two separate keyword terms, “addictive” and “behavior.”

Because the term “gamble” was the initial search term, it was distributed broadly, occurring in 71% of all citations; this distribution was too widespread to use as a basis for creating groups. Therefore, we excluded gambling and its associated synonyms (e.g., gambles, gambled, gambling) from the analysis. We also excluded the 36 two-letter abbreviations described above (e.g., “/px”) because of their redundancy with the bracketed terms (e.g., “/px” always precedes “[psychology]”).

This procedure yielded a total of 1,240 unique keyword terms distributed among 2,194 (i.e., citations that had keywords other than gambling) of the 2,246 citations. RefViz used relationships among these terms to categorize citations into 47 groups, a number determined by the square root of the number of citations. More specifically, RefViz uses a word-based statistical analysis to extract key concepts from the prevalence and co-occurrence of words within citations. Words that occur together more or less frequently than they would by chance are combined to create categories. The software assigns a numeric value between 1 and -1 to each term representing its association with each group and assigns each citation to the one group with which its keywords are most related. We named each category according to the keyword(s) with the highest-loading on that category. We conducted this analysis on both the full sample and the subset published between 1999 and 2003.

Results

Frequency Analysis

Although there is substantial variation in the number of publications per year, evidence reveals an increasing number of gambling-related citations over time. Covering 100 years of gambling research, curve estimation revealed that an exponential curve best fit the relationship between year of publication (i.e., 1903–2003) and publication frequency ($n = 2,246$), accounting for 83% of the variance in number of articles per year, $F(1, 72) = 361.70$, $P < .001$. The vast majority (97%) of publications were released after 1963; Figure 1 illustrates the distribution of this subset of gambling-related references stratified by year of publication (i.e., from 1964 to 2003, $n = 2,176$). As observed with the larger set of citations, this subset similarly reveals an exponential growth curve, $F(1, 38) = 89.55$, $P < .001$, that accounted for 70% of the variance associated with publication frequency.

To confirm whether the increasing trend that Eber and Shaffer (2000) had previously identified continued beyond their last observation, we limited the next analysis to data collected subsequent to 1998. Regressing publication frequency on the target years 1999–2002 ($n = 596$) revealed a statistically significant positive linear trend, $F(1, 2) = 232.32$, $P < .01$, that accounted for 99% of the variance associated with publication frequency. Gambling-related studies published since 1998 (i.e., between 1999 and 2003) disproportionately contain about 1/3 of all gambling-related citations in this sample.

Categorical Analysis by Keyword

Figure 2 summarizes the most frequent primary keywords in references from 1903 to 2003 that appeared in 2% or more of the 2,246 citations. Although not illustrated in Fig. 2 because their inclusion would obscure the other results, the keywords “gambling” and “pathological gambling” were the most common keywords, occurring in 46% and 28% of citations, respectively. As Fig. 2 shows, risk-taking, which occurred in 9% of citations, and decision-making, which occurred in 7% of citations were the next most common keywords, followed by addiction and choice behavior, each of which occurred in 3% of the citations. The 1,648 keywords

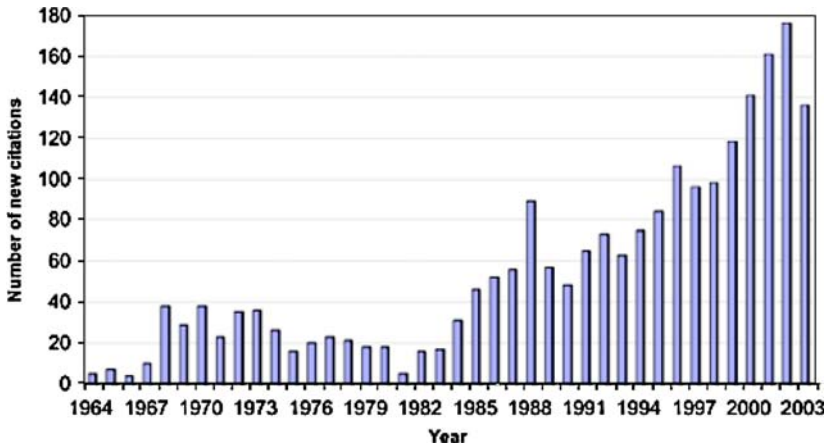


Fig. 1 Frequency of citations from 1964 to 2003

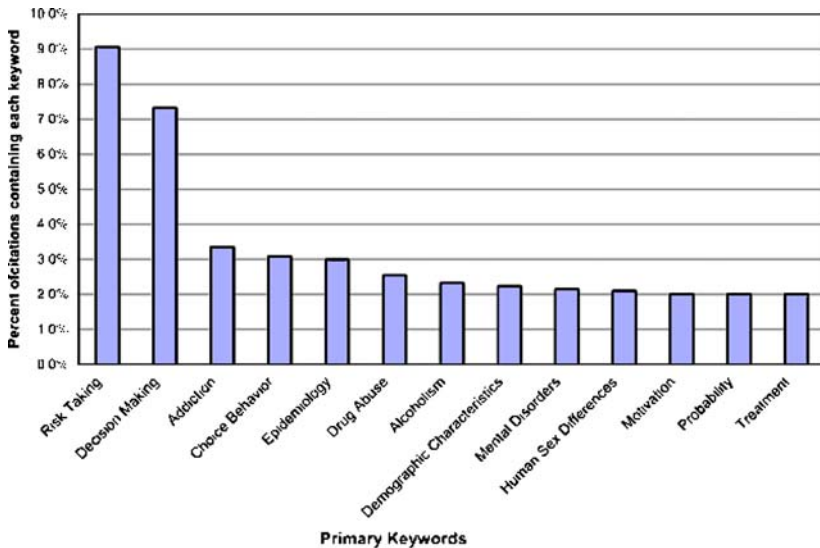


Fig. 2 Prevalence of primary keywords in gambling citations between 1964 and 2003. Note. “Gambling” and “pathological gambling,” though not included in this figure, were the most prevalent keywords, occurring in 46% and 28% of citations, respectively. Ninety-two percent of citations also made reference to one of the keywords represented by the “Other” category. The “Other” category is an aggregate of keywords, each of which appeared in fewer than 2% of citations

included in the “other” category occurred in 92% of citations; 17% of the citations referred only to these “other” keywords.

Figure 3 summarizes the most cited topics of interest in citations published between 1999 and 2003 ($n = 732$). To avoid obscuring the other results “gambling” and “pathological gambling” were again the most common keywords (contained by 44% and 35% of citations, respectively) and were again excluded from display. In this subset of recent citations, as compared to the overall sample, decision-making surpassed risk-taking to become the third most prevalent topic (behind “gambling”

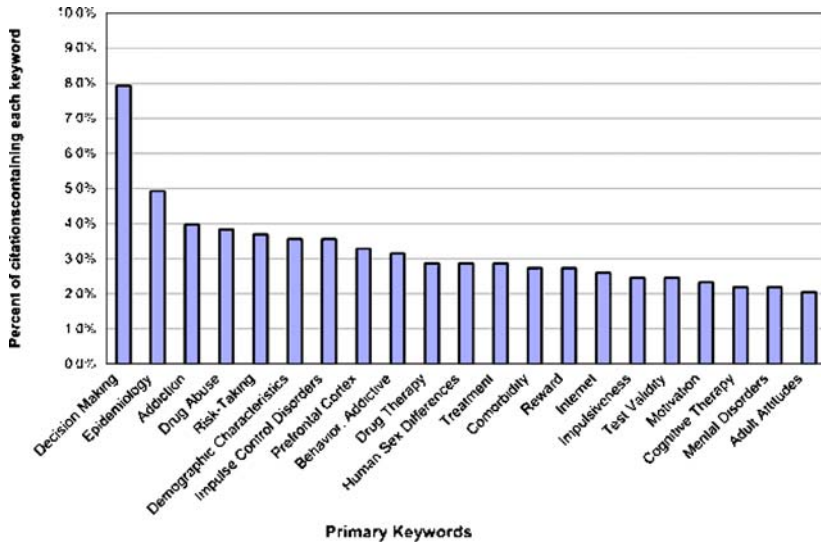


Fig. 3 Prevalence of primary keywords in gambling citations between 1999 and 2003. *Note.* “Gambling” and “pathological gambling,” though not included in this figure, were the most prevalent keywords, occurring in 44% and 35% of citations, respectively. Ninety-one percent of citations also made reference to one of the keywords represented by the “Other” category. The “Other” category is an aggregate of keywords, each of which appeared in fewer than 2% of citations

and “pathological gambling”) in gambling-related journal articles. Reference to epidemiology also increased substantially, surpassing choice behavior and addiction to become the fifth most prevalent primary keyword. Other notable differences from the complete dataset include increased prevalence of citations that focus on drug abuse, addiction, demographic characteristics, impulse control disorders, comorbidity, and neuroscience (e.g., “prefrontal cortex,” “drug therapy”). The 709 keywords included in the “other” category occurred in 91% of citations, 10% of which referred to these keywords exclusively.

Categorical Analysis of Keywords Using RefViz

By applying the RefViz clustering analysis to all keyword terms (other than gambling) associated with the citations, we obtained 47 groups among which 2,194 (i.e., citations that had keywords other than gambling) of the 2,246 citations were distributed. Table 1 summarizes these groups, their highest loading keyword terms, and the number of citations associated with each group. The pure “pathological” group (i.e., the pathological group with no other highly loading keywords) contained the greatest percentage (20%) of citations. The “decision” “making” group and the “behavior” group contained 7% of citations each. Characterized by stronger negative associations with “pathologic” and “behavior” than positive associations with other keywords, the “non” psychopathology group also contained 7% of citations. “Risk-taking” defined one of the 10 most prevalent groups (containing 5% of citations), and “pathologic” had an additional strong positive association with a group (containing 6% of citations) on which “drug” loaded highly. However, “pathologic” also was associated *negatively* with several prevalent groups; seven of

Table 1 Groups and highest loading keywords extracted from RefViz analysis of 1903–2003 citation primary keywords

Group	1st term (loading)	2nd term (loading)	3rd term (loading)	# of citations	Percentage of all citations (%)
1	Pathologic (.72)	Disorder (-.10)	Psychology (-.09)	447	19.9
2	Behavior (.85)	Pathologic (-.28)	Addictive (.14)	163	7.3
3	Pathologic (-.28)	Behavior (-.14)	Disorder (-.10)	158	7.0
4	Make (.91)	Decision (.84)	Pathologic (-.27)	156	6.9
5	Psychology (.88)	Pathologic (-.28)	Experimental (.15)	145	6.5
6	Pathologic (.65)	Drug (.52)	Therapy (.50)	126	5.6
7	Risk-taking (.90)	Pathologic (-.28)	Behavior (-.14)	105	4.7
8	Choice (.85)	Behavior (.79)	Probability (.30)	71	3.2
9	Disorder (.87)	Pathologic (-.28)	Physical/Psychological (.26)	69	3.1
10	Attitude (.93)	Adult (.29)	Pathologic (-.28)	58	2.6
11	Drug (.79)	Pathologic (-.26)	Effect (.23)	57	2.5
12	Disorder (.84)	Therapy (.36)	Mental (.35)	55	2.4
13	Disorder (.84)	Pathologic (.66)	Comorbidity (.19)	53	2.4
14	Health (.81)	Pathologic (-.27)	Mental/Service (.23)	48	2.1
15	Disorder (.80)	Psychology (.46)	Behavior (.33)	46	2.0
16	Human (.91)	Sex (.66)	Difference (.66)	32	1.4
17	Personality (.91)	Pathologic (-.28)	Correlate (.25)	32	1.4
18	Probability (.78)	Judgment (.56)	Pathologic (-.28)	32	1.4
19	Control (.87)	External (.80)	Internal/Locus (.77)	30	1.3
20	Social (.90)	Pathologic (-.27)	Ethics/Influence (.20)	30	1.3
21	Self (.76)	Perception (.34)	Pathologic (-.28)	29	1.3
22	Addiction (.96)	Pathologic (-.26)	Internet (.13) Behavior (-.13)		231.0
23	Test (.90)	Validity (.50)	Reliability (.25)		200.9
24	Game (1.0)	Computer (.50)	Delinquency/Juvenile (.25) Pathologic (-.25)		160.7
25	Theory (.81)	Psychoanalytic (.25)	Pathologic (-.25)		160.7
26	Characteristic (.80)	Demographic (.47)	Job (.47)		150.7
27	Electric (.79)	Countershock (.71)	Heart (.64)		140.6
28	Relation (.43)	Alcoholism (.29)	Pathologic (-.29)		140.6
29	Genetics 1 0	Dopamine (.67)	Receptor (.67)		120.5
30	Psychotherapy (.67)	Group (.50)	Process (.33)		120.5
31	Treatment (.58)	Personnel (.42)	Pathologic (-.25)		120.5
32	Suicide (.58)	Internet (.33)	Factor (.25) Pathologic (-.25)		120.5
33	Difference (.91)	Age (.55)	Sex/Human (.27) Pathologic (-.27)		110.5
34	Reinforcement (.90)	Schedule (.60)	Pathologic (-.30)		100.4
35	Compulsion (.50)	Emotional (.50)	State (.40)		100.4
36	Physiologic (1.0)	Arousal (.88)	Correlate/Heart/Rate/Sensation/Seek (.25) Pathologic (-.25)		80.4
37	Statistics (.88)	Data (.38)	Numerical/Psychometrics (.38)		80.4
38	Evoke (.86)	Potential (.86)	Money (.57)		70.3
39	Depression (.86)	Major (.43)	Emotion (.29) Pathologic (-.29)		70.3
40	Criminals (.33)	Pathologic (-.33)	Injury/Etiology (.33)		60.3
41	Expectation (1.0)	Prediction (.25)	Prisoner/Strategy (.25) Pathologic/Behavior (-.25)		40.2

Table 1 continued

Group	1st term (loading)	2nd term (loading)	3rd term (loading)	# of citations	Percentage of all citations (%)
42	Development (1.0)	Pathologic (-.25)	Catholicism/Child/Play/ Childhood/Playthings/ Moral/Protestantism/ Psychosocial/Roman/ Factor/Adolescent (.25) Behavior (-.25)		40.2
43	Interpretation (1.0)	Psychoanalytic (1.0)	Literature (.67)		30.1
44	Delinquency (1.0)	Juvenile (1.0)	Crime (.67)		30.1
45	Design (.50)	Pathologic (-.50)	Evaluation/Interior/ Esthetic (.50)		20.1
46	Attribution (1.0)	Pathologic (-.50)	Cycle/Expression/Facial/ Failure/Tension/Premen- strual/Menstrual (.50)		20.1
47	Hypersexuality (1.0)	Literature (1.0)	Review (1.0)		10.0

Note. Bolded terms are the most prevalent keyword terms in each group. Shaded area includes groups that *cumulatively* account for less than 10% of the citations in the sample

the 10 most prevalent groups included this negative association as one of their top three keyword associations. The cluster analysis reveals categorical representation (i.e., classification into one of the 47 groups) for 98% of all the citations published between 1903 and 2003.

When we conducted the same RefViz analysis as before using only citations published from 1999 to 2003 ($n = 732$), the clustering procedure produced 27 groups among which 716 of the 732 citations were distributed. Table 2 shows that the group containing the greatest percentage of citations (28% of citations) was again most strongly associated with the term “pathologic.” The terms “decision” and “make,” associated in the full sample with the third most prevalent group, helped define the second most prevalent group (14% of citations) in this subset. A group defined by the terms “drug,” “abuse,” and “therapy” (12% of citations) rounded out the top three most prevalent groups, replacing the full sample’s “behavior” group in the top three. For the recent subset of citations, “risk-taking,” which loaded heavily on a prevalent group obtained from the full set of citations, was not strongly associated with the most prevalent groups. “Pathologic” was again negatively associated with several prevalent groups; five of the ten most prevalent groups included this negative association with “pathologic” as one of their top three keyword associations. The cluster analysis of this subset of data reveals categories for 98% of all the citations published between 1999 and 2003.

Discussion

The growth of psychological and biomedical publications about gambling represents a significant increase in gambling-related research. Pathology, risk-taking, decision-making, and addiction have dominated this body of research during the 20th century.

Table 2 Groups and highest loading keywords extracted from RefViz analysis of 1999–2003 citation primary keywords

Group	1st term (loading)	2nd term (loading)	3rd term (loading)	# of citations	Percentage of all citations (%)
1	Pathologic (.63)	Psychology (–.12)	Drug (–.12)	204	27.9
2	Decision (.39)	Make (.39)	Pathologic (–.33)	101	13.8
3	Drug (.78)	Abuse (.28)	Therapy (.26)	85	11.6
4	Behavior (.76)	Pathologic (–.22)	Addictive (.22)	63	8.6
5	Disorder (.78)	Control (.45)	Impulse (.37)	51	7.0
6	Psychology (.76)	Pathologic (–.35)	Drug (–.12)	51	7.0
7	Attitude (.67)	Pathologic (–.36)	Adult (.30)	33	4.5
8	Health (.78)	Pathologic (–.33)	Public/Smoke (.22)	18	2.5
9	Test (.91)	Validity (.91)	Reliability (.45)	11	1.5
10	Difference (.80)	Human (.70)	Sex (.70)	10	1.4
11	Motivation (.60)	Pathologic (–.40)	Experimental (–.30)	10	1.4
12	Electric (1.0)	Countershock (1.0)	Therapy/Heart/Arrest (.67)	9	1.2
13	Adverse (.89)	Effect (.89)	Levodopa/Agents (.56)	9	1.2
14	Characteristic (.88)	Demographic (.63)	Pathologic (–.38)	8	1.1
15	Probability (1.0)	Judgment (.57)	Statistical (.43)	7	.9
16	Genetics (1.0)	Polymorphism (.67)	Monoamine/Disease/D2/Carrier/Receptor/Glycoprotein/Protein/Membrane/Oxidase/Dopamine (.33)	6	.8
17	Analysis (1.0)	Society (.50)	Cost (.50)	6	.8
18	Frontal (.83)	Lobe (.83)	Money (.67)	6	.8
19	Physiologic (.83)	Arousal (.67)	Heart/Rate (.67)	6	.8
20	Internet (1.0)	Pathologic (–.33)	Addiction/Clinician/Compulsion/Schema/Counsel/Counselor/Psychosocial/Myth/Diagnostic/Research/Criterion/Factor (.17) Behavior/Disorder/Drug/Make/Therapy/Psychology (–.17)	6	.8
21	Human (.80)	Sensation (.60)	Seek (.60)	5	.7
22	Internal (1.0)	External (1.0)	Locus/Control (1.0)	3	.4
23	Alcohol (1.0)	Drink (1.0)	Pattern (1.0)	3	.4
24	Relation (1.0)	Pathologic (–.50)	Parent-Child/Physician-Patient/Parent (.50)	1	.1
25	Public (1.0)	Medium (1.0)	Mass/Opinion (1.0)	1	.1
26	Theory (1.0)	Formulation (1.0)	Chance/Fortune (1.0)	1	.1
27	Prediction (1.0)	Extrasensory (1.0)	Perception (1.0)	1	.1

Note. Bolded terms are the most prevalent keyword terms in each group; Shaded area includes groups that cumulatively account for less than 10% of the citations in the sample

Citation data reflect the progress of gambling-related research and confirm a previously identified increasing trend. More specifically, 97% of gambling-related articles have been published after 1963; further, publications released between 1999 and 2003 disproportionately contain almost 1/3 of all gambling-related citations in this sample. The new data and research methods in this study suggest a potentially different interpretation of the subject matter included in these citations than Eber and Shaffer (2000) offered previously.

Since 1903, scientific gambling-related citations have been increasing at an exponential rate. Using new methods, the results of the current study confirm the earlier observations of Eber and Shaffer (2000), whose study included citations sampled from 1964 to October 22, 1999. The new results also identify a continuing positive growth trend from 1999 to 2002.

By more comprehensively categorizing citations by topic, this study yielded a different but related set of findings to those Eber and Shaffer (2000) obtained previously. As before, this study examined those topics most frequently found in citations. Unlike the previous research, in one analysis, we allowed each citation to refer to multiple primary categories, and in another analysis we used objective mathematical techniques to cluster citations by category. Eber and Shaffer (2000) found that the most prevalent topic of gambling studies was cognition or personality followed by methods or theory. Using the first technique, this study, however, revealed that the most prevalent topics of gambling research after “gambling” were “pathological gambling,” “risk-taking,” “decision-making,” and “addiction.” Similarly, using the RefViz mathematical clustering procedure, we identified that the most prevalent clusters included terms referring to pathology, decision-making, and risk-taking.

The concept of psychopathology has been a vital element within gambling studies during the last 100 years, as reflected by the frequency of the keyword “pathological gambling.” Although Eber and Shaffer (2000) noted that cognition and epidemiology were important topics of gambling studies, the current analysis implies that it is also important to consider the place of pathology-related studies in gambling research. The fact that reference to pathology is more prevalent in citations from 1999 to 2003 than among those published during the last 100 years suggests increasing attention to the relationship between psychopathology and excessive gambling. This result likely reflects the addition of pathological gambling to the nomenclature of the American Psychiatric Association’s diagnostic manual in 1980. However, as Tables 1 and 2 show, many areas of gambling studies are not associated with pathology. More than half of the clusters in both the full set and subset of citations have strong negative associations with pathology. As evidenced by these clusters that are defined as much by a negative association or absence of reference to pathology as by their inclusion of other keywords, gambling research seems to be bifurcating between (a) gambling studies and (b) pathological gambling studies. This divergence might reflect an increasingly multidimensional understanding of various types of gambling behavior, reflecting a growing appreciation for the complexities found among various patterns of gambling (e.g., legal/illegal, recreational, excessive, intermittently intemperate, etc.).

Both categorical analyses in this study highlight the importance of cognition and decision-making as foci of gambling-related research. While the pathological aspects of gambling behavior are prevalent in the dataset, many citations also focus on research that features decision-making and cognitive processes as areas of inquiry. In fact, the clustering analysis revealed that the most prevalent group on which decision-making loaded highly also was defined by its *lack* of association with pathology (i.e., pathologic loaded negatively on this group—see Tables 1 and 2).

The growing research focus on the internal mechanisms associated with gambling has evolved recently to incorporate advances in the field of neuroscience. The primary keyword data collected from the citations between 1999 and 2003 reveals that “prefrontal cortex” and “drug therapy” (i.e., treating mental disorders with drugs

designed to affect specific areas of the brain) have become two of the most prevalent keywords listed by citations published during these 5 years. Similarly among the 27 clusters created from this subset of data, “drug” and “therapy” load highly on a group that represents 12% of the citations in this sample.

Other changes have been taking place in the field of gambling-related research. In the past 5 years (i.e., 1999–2003), “epidemiology” has replaced “risk-taking” as one of the three most prevalent keywords, impulse control disorders have entered the top ten, and drug abuse, comorbidity, and demographic characteristics have become more prominent topics. These changes reflect a trend toward placing gambling in a larger context—understanding not only the behavior, but also recognizing its correlates, antecedents, and consequences. In the study of epidemiology, research often moves from descriptive study of a phenomenon and its distribution to a causal investigation of the determinants (i.e., risk and protective factors) that increase or decrease the likelihood of the phenomenon developing (Regier & Robins, 1991). The study of gambling appears to have developed similarly, moving from research that describes the behavior to investigations of its context and determinants (Shaffer & Kidman, 2004; Shaffer, LaBrie, LaPlante, Nelson, & Stanton, 2004).

The use of two different analytic techniques for detecting key concepts in the gambling literature provides the benefit of confirming the reliability of findings that co-occur across method, but also raises questions about findings that do not coincide from analysis to analysis. In this case, the overall results corresponded well between the two techniques—pathology, decision-making, and risk-taking emerged as the most important concepts for both analyses. However, RefViz’s analysis corresponded only moderately with the analysis of keywords for the key concepts of the last 5 years of citations. “Decision-making” and “epidemiology” did not load above .5 on any of the ten most prevalent clusters for the citations published between 1999 and 2003, even though they were among the most prevalent keywords during those 5 years.

These two techniques (i.e., the analysis by keyword and the clustering procedure) provide different perspectives that might account for this discrepancy. The straightforward keyword analysis considers the keywords as they appear within each citation and obtains a count of their prevalence. The clustering analysis moves beyond the independent representation of each keyword and identifies the relationships *between* keywords in addition to the prevalence of keywords within citations. As a result, keywords that are highly prevalent in citations but do not co-occur regularly with other keywords will not load strongly on the clusters RefViz creates (e.g., 1999–2003 references to “epidemiology”). However, the clustering procedure identified the common co-occurrence of words such as “pathological” and “drug” within citations published between 1903 and 2003, a relationship not recognized by the keyword analysis. Therefore, we suggest that the apparent discrepancies between these two techniques are more accurately viewed as complementary perspectives than contradictions. Together, the two techniques offer more information than either could provide alone.

Given the subtle but important differences in category-based results between the current and the earlier gambling citation study, the results of this study suggest that assigning each citation to only one category was insufficient to fully capture the multidimensional subject matter of gambling-related research. By using multiple and different dependent variables that purport to measure the same underlying construct, research can achieve more valid measurement of these constructs

(Campbell & Fiske, 1959). Within the keyword analysis, we used multiple categorical measurements of each citation to ensure the validity of the underlying classification. Where Eber and Shaffer (2000) used only one data point per citation to classify 1,181 publications, this analysis used approximately three data points per citation (i.e., 6,165 data points) to classify 2,246 publications, providing a more detailed and inclusive analysis of citation topics. The single data point strategy left 21% of the citations as “unclassifiable”; the multiple data point strategy, however, yielded less than 1% of citations that were unclassifiable and only 17% of citations that did not fit into our top 13 primary keyword categories. Our clustering procedure successfully classified 98% of citations. Although the earlier study, which reduced each citation to one topic, could explore the trends of only the primary subject of each article, by identifying multiple keywords within each citation, this new strategy provides additional information about the range and frequency of topics targeted by gambling-related research. By using multiple methods in addition to multiple measurements and relying on objective criteria for detecting key concepts in the citation literature, this study ensures that the findings are robust and replicable.

Caveats

The current study’s reliance on keywords entails three primary caveats. First, authors have used keywords increasingly over time. No citations prior to 1963 had more than two keywords of any kind. Between 1999 and 2003, citations averaged more than three *primary* keywords, averaged more than eight keywords overall, and contained up to 54 keywords. Thus, changes in keyword patterns across time might reflect the increase in overall use of keywords, not just specific topics.

Second, both categorical analyses utilized in this study were essentially linguistic analyses. Although keywords represent as good a measure as any of underlying content, by relying on the words and not their underlying meaning to represent categories, we might have missed important conceptual relationships. For example, the keywords “dopamine,” “serotonin,” “neuroscience,” and “prefrontal cortex” all refer to the underlying domain of neuroscience, but are kept separate for both analyses unless they co-occur within the same citation.

Finally, in addition to the analytic caveats, there are two limitations associated with research that relies on the major databases of published work to identify a field’s research trends. First, by relying on databases, citation research might not represent articles published in journals not contained in those databases (e.g., economic journals not included by PsychInfo or MedLine). Second, the research cannot depict the most current patterns of research in the field. By time a study is published, because of the review time and the publication lag, it represents research that is at least one and possibly 2 years old.

Recommendations

This study reveals that some topical areas in gambling studies have been more thoroughly investigated than others (e.g., pathological gambling and decision-making). This disproportionate attention to certain themes suggests that some areas need more attention. Neuroscience, genetic, and drug trial studies have only recently been observed in biomedical and behavioral journals. In coming years, it will be important

to encourage this type of research. The information gained from these types of studies will add much needed knowledge to complement the already flourishing psychological and behavioral research in gambling studies publications. As stated earlier, gambling research is beginning to study not only gambling behavior and its mechanisms, but the context within which it occurs. Together, these varied elements can better inform the treatment of gambling problems and contribute to the goals of a public health model of gambling: (a) learning more about gambling decisions and behavior, (b) identifying those determinants of disordered gambling, and (c) providing a range of interventions for gambling-related problems (Korn & Shaffer, 1999; Shaffer & Korn, 2002).

This study also suggests that there is a need for procedural standards in citations management. With the emergence of digital methods for searching keywords to identify new and old research alike, it becomes increasingly important to provide an easy, unambiguous and reliable method to review the extant literature. Standards used to facilitate this process might include a greater use of the commonly prescribed keywords terms, database searches in standard fields (i.e., keyword section and not title field), and the use of only a few particular search engines with standardized interfaces (i.e., PsycINFO/MEDLINE or Inspec/Alt HealthWatch). In the future, these changes will allow for a standardization of search inquiries; such consistency holds the potential to yield increasingly more reliable and replicable literature searches, as well as more informative descriptions of the trends present in both nascent and mature academic fields.

Acknowledgments The authors extend special thanks to Gabriel Caro, Tony Donato, Debi LaPlante, and Christine Reilly for their important contributions to this project.

References

- Adair, J. G., & Vohra, N. (2003). The explosion of knowledge, references, and citations. Psychology's unique response to a crisis. *American Psychologist*, 58(1), 15–23.
- American Psychological Association (2004). PsycINFO®. Retrieved January 20, from <http://www.ovid.com/site/catalog/DataBase/139.jsp?top = 2&mid = 3&bottom = 7&subsection = 10>.
- Brown, A. J., & Richman, D. D. (1997). HIV-1: Gambling on the evolution of drug resistance? *Nature Medicine*, 3(3), 268–271.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Eber, G. B., & Shaffer, H. J. (2000). Trends in bio-behavioral gambling studies research: Quantifying citations. *Journal of Gambling Studies*, 16(4), 461–467.
- Garfield, E. (1955). Citation indexes for science; a new dimension in documentation through association of ideas. *Science*, 122, 108–111.
- Henderson, A. (1998). Electronic databases and publishing. New Brunswick, N.J.: Transaction Publishers.
- Korn, D. A., & Shaffer, H. J. (1999). Gambling and the health of the public: Adopting a public health perspective. *Journal of Gambling Studies*, 15(4), 289–365.
- Microsoft Corporation. (2002). Microsoft Excel 2002 (Version 10 SP-2). Redmond, WA: Microsoft Corporation.
- Najman, J., & Hewitt, B. (2003). The validity of publication and citation counts for sociology and other selected disciplines. *Journal of Sociology*, 39(1), 62–80.
- National Gambling Impact Study Commission (1999). *National gambling impact study commission report*. Washington, DC: National Gambling Impact Study Commission.
- National Research Council (1999). *Pathological gambling: A critical review*. Washington DC: National Academy Press.
- OID Technologies (2003). OVID (Version 9.0). New York City: OVID Technologies.

- Rait, C. (1989). The AMA: Gambling with nursing's future. *Neonatal Network—Journal of Neonatal Nursing*, 7(5), 7.
- Regier, D. A., & Robins, L. N. (1991). Introduction. In D. A. Regier, & L. N. Robins (Eds.), *Psychiatric disorders in America: The epidemiologic catchment area study* (pp. 1–10). New York: The Free Press.
- Shaffer, H. J., & Kidman, R. C. (2004). Gambling and the public health. In J. E. Grant & M. N. Potenza (Eds.), *Pathological gambling: A clinical guide to treatment* (pp. 3–23). Washington, DC: American Psychiatric Publishing, Inc.
- Shaffer, H. J., & Korn, D. A. (2002). Gambling and related mental disorders: A public health analysis. *Annual Review of Public Health*, 23, 171–212.
- Shaffer, H. J., LaBrie, R. A., LaPlante, D. A., Nelson, S. E., & Stanton, M. V. (2004). The road less travelled: Moving from distribution to determinants in the study of gambling epidemiology. *Canadian Journal of Psychiatry*, 49, 504–516.
- SPSS Inc. (2003). SPSS (Version 12.0) [statistics program]. Chicago: SPSS.
- Thomson ISI ResearchSoft (2003). EndNote (Version 7.0) [World Wide Web]: Thomson ISI ResearchSoft.
- Thomson ISI ResearchSoft (2004). RefViz (Version 1.0) [citation searching program].
- Thorngate, W. (1990). The economy of attention and the development of psychology. *Canadian Psychology*, 31(3), 262–271.
- U.S. National Library of Medicine (2004). MEDLINE®. Retrieved January 20, from <http://www.ovid.com/site/catalog/DataBase/900.jsp?top = 2&mid = 3&bottom = 7&subsection = 10>.
- Xhingnesse, L., & Osgood, C. (1967). Bibliographic citation characteristics of the psychological journal network in 1950 and in 1960. *American Psychologist*, 22, 778–791.