Technological Forecasting & Social Change 146 (2019) 181–202

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Evaluating university industry collaborative research centers

Elizabeth Gibson[a], Tugrul U. Daim[b,c,*], Marina Dabic[d,e]

[a] University of Colorado, Boulder, CO, USA
[b] Portland State University, Portland, OR, USA
[c] Chaoyang University of Technology, Taichung, Taiwan
[d] Nottingham Trent University, UK
[e] University of Zagreb, Croatia

ARTICLE INFO

Keywords:
Research and Development Management
Hierarchical Decision Modeling
Industry University Collaboration
National Science Foundation

ABSTRACT

This research provides performance metrics for cooperative research centers that enhance translational research through partnerships formed by government, industry and academia. Centers are part of complex ecosystems and vary greatly in the type of science conducted, organizational structures and expected outcomes. The ability to realize their objectives depends on transparent measurement systems to assist in decision making in research translation. We introduce a hierarchical decision model that uses both quantitative and qualitative metrics. A generalizable model is developed based upon program goals. The results are validated through consultation with experts. The method is illustrated using data from the National Science Foundation's industry/university co-operative research center (IUCRC) program. The methodology provides a basis for a generalizable model and measurement system to compares performance of university science and engineering focused research centers supported by industry and government.

1. Introduction

Industry-university collaborations conducting multi-disciplinary research are required to solve increasingly complex social problems (Boardman and Gray, 2010). Increased U.S. public policy support for initiatives that enhance translational research has resulted in the evolution of many different forms of technology transfer mechanisms (Boardman and Bozeman, 2015). Today, university-based research centers "are prevalent as both policy mechanisms and industry strategies" [(Boardman and Ponomariov, 2011) pg 76]. Cooperative research centers (CRCs) that involve partnership agreements with actors from three different sectors of government, academia and industry are the most sustainable business models (Lee, 2000). However, supporting these "triple-helix" (Etzkowitz and Leydesdorff, 2000a) or government-university-industry (GUI) (Carayannis et al., 2014a) collaborations is expensive, driving policy makers to shift their attention towards performance evaluation.

Academia, policy makers (Perkmann et al., 2011a) and CRC managers are all invested in understanding the performance and impact of these centers (Bozeman et al., 2013a). A wealth of literature examines program evaluation through primarily qualitative case-based methods or quantitative methods based on traditional indicators such as patents and publications. Despite the effort and many excellent studies, researchers are cautioning that traditional measures are inadequate (Gray et al., 2014a), placing a call-to-arms for further research. A multi-dimensional-holistic study with a flexible approach that can evaluate both quantitative and qualitative output indicators is needed. This research begins to fill this gap by presenting a generalizable model for CRC performance evaluation.

The National Science Foundation (NSF) is responsible for technology planning and science and engineering based research and education in the United States. Recognizing the value of industry sponsored cooperative research, the NSF launched a program in 1980 to improve the linkage between industry and university for cooperative research (Gray et al., 2012a); now known as the Industry-University Cooperative Research Center (IUCRC) program. The success of this model led to the development of other NSF science and engineering centers. Because the model has been replicated multiple times, the social technology clarifies the unit of analysis making it a better candidate for study than other CRCs. Today, over 66 IUCRCs are actively supported by the NSF. Literature shows the IUCRC to be one of the more successful CRCs (Geisler, 2003).

Supporting such centers is expensive. So, academia, policy makers and (Perkmann et al., 2011b) CRC managers are all invested in understanding the performance and impact of these centers (Bozeman et al., 2013b). Researchers acknowledge that "the growth in private and

* Corresponding author at: Portland State University, Portland, OR, USA.
E-mail address: tugrul.u.daim@pdx.edu (T.U. Daim).

https://doi.org/10.1016/j.techfore.2019.05.014
Received 16 April 2018; Received in revised form 24 April 2019; Accepted 4 May 2019
Available online 07 June 2019
0040-1625/

public investment in university-based technology initiatives has raised important policy questions regarding the impact of such activities (Link and Siegel, 2005a; Phan and Siegel, 2006). This interest has led to a wealth of literature examining program evaluation through primarily qualitative case-based methods or quantitative methods based on traditional indicators such as patents and publications.

Despite the effort and many excellent studies, researchers are cautioning that traditional measures are inadequate (Gray et al., 2014b), placing a call-to-arms for further research.

This study examines the literature to explore the concerns about current indicators and measurement systems. It adds value by developing a flexible measurement system incorporating qualitative and quantitative metrics. A generalizable model is developed that uses a holistic and balanced approach to produce a score that measures effectiveness in which a center is achieving the NSF program's mission. The Wood Based Composites Center illustrates the method using actual center data. Experts validated the methodology and results adding a transparent decision support tool for performance evaluation into the stream of literature.

Including this introduction, the paper is organized into 6 sections. Section 2 reviews the academic literature on national planning of technology and cooperative research center program evaluation. Section 3 describes the research approach and methodology. Section 4 shows how a generalizable hierarchical decision model (HDM) is developed and finalized using expert judgment. Section 5 illustrates and validates the method using actual data collected for a selected IUCRC, discusses the results and summarizes the expert response to the criterion related validation. Finally, Section 6 concludes the paper.

2. Literature review

2.1. National technology planning

Societal goals change throughout time driving national planning activities and the creation of public policies. Technology foresight is a process that systematically looks into the future to examine areas of research and emerging technologies that can help address changing societal goals. Technology foresight has also been defined as a tool in policy and strategic planning to "wire-up" national innovation programs (Grupp and Linstone, 1999), for priority setting and decision making (Ecken et al., 2011) and for creation of vision and the pursuit of knowledge (Yokoo and Okuwada, 2013).

Public policy strategies are often the outcome of national foresight activities (Georghiou and Cassingena Harper, 2013). Previous to 1980, US policies traditionally focused on facilitating collaboration among industry and academia (Gibson, 2015; Martin and Johnston, 1999); then the national research agenda shifted to place more focus on technology transfer. Initiatives to facilitate technology transfer have been developed using a variety of different mechanisms that vary in terms of complexity, structure and longevity including: research parks, licensing agreements, R&D limited partnerships, joint facility use agreements, research institutes, research centers and state-supported science and technology centers. The most sustainable technology transfer mechanisms require industry-sponsored collaborative research (Link and Siegel, 2005a).

"System changes are labelled 'socio-technical' because they not only entail new technologies, but also changes in markets, user practices, policy and cultural meanings"(Geels, 2010). Major industries such as information and communication technology (ICT)(Rohrbeck, 2010), energy(Rohrbeck et al., 2013), food(Chavez, 2013), health (Masum et al., 2010) and transportation (Alkemade and Suurs, 2012) are faced with complex socio-technical challenges. Solving environmental problems is a national concern that entails cultural value and belief systems [29]that goes far beyond a technological problem.

2.2. Technology research centers

Roessner defines technology transfer as "the movement of know-how, technical knowledge, or technology from one organizational setting to another"[(Roessner, 1998) p 31]. The university ecosystem began to change to support technology transfer as evidenced by the creation of technology transfer offices (TTOs) (Link and Siegel, 2005b) and mission expansion to include entrepreneurial and commercialization statements (Tran, 2013a). Interested in further supportive policies, government started looking for practical organizational structures (Daim et al., 2006; Robinson et al., 2013) that encouraged knowledge and technology transfer (Carayannis and von Zedtwitz, 2005) beyond the university sector (Smilor et al., 1989).

Studies provide evidence that public funding of research has had significant impact on CRC programs (Cunningham et al., 2014) recording over 27,500 different CRC programs worldwide and thousands in the US alone. A variety of different mechanisms developed that vary in terms of complexity, structure and longevity including: research parks, licensing agreements, R&D limited partnerships, joint facility use agreements, research institutes, research centers and state-supported science and technology centers.

Fig. 1 shows how three of the NSF sponsored CRC programs are positioned in the middle level of performance evaluation problems: materials science and engineering research centers (MRSECs), engineering research centers (ERCs) and industry/university cooperative research centers (IUCRCs).

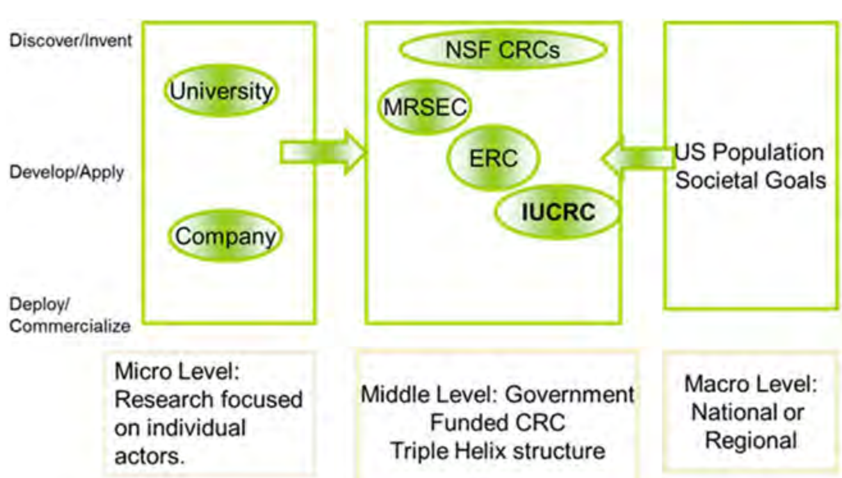


Fig. 1. CRCs are ecosystems.

Today, GUI CRC's are a popular mechanism (Geisler, 2003; Gray et al., 2012b) for translational research because industry funding is an important business model component for sustainable innovations (Gray and Walters, 1998; Rohrbeck and Kaab, 2013). Bozeman named one stream of literature the "cooperative technology policy paradigm" because it "features an active role for government actors and universities in technology development and transfer" [(Bozeman, 2000) p 632]. Experts are concerned that *"evaluating such centers remains difficult and often subjective, yet federal science agencies continue to invest considerable resources in them."* They are resource intensive and financially expensive (Gray et al., 2013) receiving over $5 billion in federal funding (Gray, 2008) for support and evaluation.

Several examples show how policy makers have responded: Passed in 1993, the Government Performance and Results Act requires codification of the use of quantitative metrics for program evaluation (Kostoff, 2005). In 2010, the America Competes Act Reauthorization was passed to further support linkages between research investments and economic growth and societal benefits (Cragin et al., 2012).

2.3. Evaluation methods

The evaluation method literature was synthesized into five (Etzkowitz and Leydesdorff, 2000b) groups for further discussion:

1) quantitative econometric and statistical analysis,
2) case-based analysis,
3) social network analysis (SNA),
4) multiple criteria decision making (MCDM),
5) multi-attribute utility theory (MAUT).

One comprehensive report by Ruegg & Feller (Ruegg and Feller, 2003) that surveyed evaluation methods and models was particularly useful. The rest of this section reviews the evaluation methods in the literature by the five research method groupings.

Licensing performance is a common theme in the quantitative based literature. For example, Chukumba and Jensen (Chukumba and Jensen, 2005), examine how the characteristics of different actors affect licensing performance. Two findings provide evidence of a positive relationship between the importance of venture capital and licensing agreements and that engineering faculty was relatively more important than the other science based faculty. Anderson et al. used licensing data to examine efficiency (Anderson et al., 2007a) and Kim took an in-depth look at the impact of lag time using similar data and metrics. Shane and Somaya(Shane and Somaya, 2007a) use the association of university technology transfer managers (AUTM) association data and patent litigation data to examine the effects on university licensing efforts (Shane and Somaya, 2007b).

The Feldman and Kelly study is different because it uses statistical analysis to test the strength of hypothesized relationships. This method is interesting because it can help to open up the "box" and take a look at the "middle". The survey data was coded as a bi-variate "yes" or "no" then tabulated and tested for statistical significance. Logistic regression was used to test the strength of the relationships (Hall et al., 2003).

The research method selected for a study depends upon the research problem being investigated and the organizational structure under analysis (Hicks, 2012). These studies are particularly useful at the micro, single-actor level, or macro, total-program level because they use a more consistent method that can be replicated by other researchers to verify and extend the results building convincing evidence about program impacts. These methods are particularly useful to justify the existence of a program and investigate if the total cost of the policy is beneficial to society. Table 1 shows some methods and findings; however, the results don't help to provide comparisons between centers.

The NSF's Industry-University Cooperative Research Center (IUCRC) program is used as the domain of organizational effectiveness for this research because based upon the longevity and formalized structure of the program. Currently there are 66 centers spanning 175 different university sites (Gray et al., 2012a; Gray et al., 2013). In the early 1980s, the NSF concerned about domestic technology transfer (Bozeman, 2000) formally launched the IUCRC program to improve the linkage between industry and university for cooperative research (Gray et al., 2012a) by transferring "know-how" in the form of organizational structure and best management practices from the NSF to a director and managing research staff.

Using a hybrid organizational structure that allows for flexibility (Gray and Walters, 1998), multidisciplinary, research projects are selected by an Industrial Advisory Board (IAB) and conducted collaboratively among university faculty, student researchers and industry partners (McGowan, 2012).

IUCRCs may take different forms and vary by participation number and levels, center goals and processes, and outputs (Hayton et al., 2010). However, there is a formal organizational model with specified policies, processes and procedures for management and evaluation. Table 2 (Gray and Walters, 1998) describes the IUCRC model by operational mechanisms and characteristics.

Formal partnership agreements are required for membership. These documents include the scope of the research projects and shared interest agreements that help to clarify intent. There are multiple stakeholders that include the NSF, the university, the center director, researchers, students and industry. Formal documents and management practices require regular reports and roadmaps. Other management practices and structural requirements help to establish an IUCRC through its' formation. For example, the funding structure requires that industrial advisory board (IAB) members pay yearly dues.

Performance appraisal is important (Abbasi et al., 2014) to the practice of CRC management to understand and maximize the impact of their research findings (Penfield et al., 2014). According to a White House memorandum (White House Office of Management and Budget, 2011), funding agencies, academic leadership, and industry must manage their portfolios in an objective, evidence-based manner to address science and technology priorities of our nation and increase the productivity of our research institutions.

The NSF has recognized the importance of a formal evaluation program by continuously supporting a project established at North Carolina State University for the purpose of evaluating IUCRCs. While the evaluation program is structured and formalized with established policies, processes and procedures to address program inputs, activities, outputs, outcomes and impacts, it struggles with some of the same challenges found in literature and is somewhat labor intensive.

In a sense, all of the NSF IUCRC program evaluators publish case studies each year for each IUCRC because they use standardized, Level of Interest and Feedback Evaluation (LIFE), forms and questionnaires to collect qualitative data. Table 3 provides an example of some of the cased-based literature focused on IUCRC performance evaluation.

Case studies are important because they can paint a detailed story and explain why events are happening tying inputs, activities, outputs and outcomes to impacts. Some of the limitations is the confidentiality of the information or the tendency to under or over report. There is also the problem of comparing centers to one another (Scott, 2014). It is difficult to generalize from a case study creating opportunity for measurement error.

Social network analysis (SNA) is gaining importance in the literature (Etzkowitz and Leydesdorff, 2000a; Perkmann et al., 2011b). Several researchers have used SNA methods, tools and techniques to investigate spill-overs (Abbasi and Altmann, 2011; Balconi and Laboranti, 2006), co-authorship networks (Abbasi et al., 2012; Li et al., 2013) and membership activity (Motoyama et al., 2014b). Structured data such as citations in the scientific databases and filings in patent databases can be mined using bibliometric techniques. Most of the researchers who use the citation of other firms' patents note that patents are not a perfect measure of innovative output (Ruegg and Feller,

Table 1
Quantitative based research in CRC literature.

Author year	Purpose	Purpose/findings	Method
Cohen et al. (1994)	Provide a comprehensive picture of IURCs	Measurement of IURC impact on technology advance	Extensive survey-based empirical study forming the "Carnegie Mellon" database
Cohen et al. (2002)	University and government research lab contribution to industrial innovation	System of simultaneous equations links dependent variables to firm/industry level economic variables	Survey-based approach using Carnegie Mellon data (1994) hypothesis-based testing
Hall et al. (2003)	Investigating roles and effects of universities in ATP-funded projects	University involvement may not speed up commercialization as partnerships may have more basic research aspects.	Survey-based study of ATP-funded research projects. Multivariate regression analysis
Chukumba and Jensen (2005)	Licensing performance focused at small business	Licensing by universities with larger venture capital, engineering faculty relative high importance	Empirical, Game theoretic model, hypothesis testing
Feldman and Kelley (2006a)	Knowledge spillover	Testing hypothesis for incentive effects of government R&D funding for firms	Empirical survey, 240 completed, multivariate regression

Table 2
IUCRC characteristics.

Characteristics	Description
Formal membership agreement	Includes unique scope and shared interest agreements
Partners	University, industry, other organizations
Shared research agenda	Objectives, goals and a roadmap
Shared IP	Formal agreement
Center director	Tied to a University (Gray and Walters, 1998), diverse (NSF, 2013)
Primary funding source	Industry members structured min. funding: $30 k from 10
Evaluation	2x/year reporting, independent formal evaluation
Graduate students	Required involvement
Structural requirements	Funding, organizational, management, reporting

2003), because they relate only to codified knowledge and there may be significant differences in patenting behavior between IUCRCs, firms, and technological domains.

However, this method shows promise and researchers are actively working to improve the problem of data availability and linkages. For example, Rafols et al. introduced a new method using bibilometric data to map areas of collaboration using network analysis methods (Rafols et al., 2009). Advances in scientific databases now allow for more sophisticated mapping and the spatial and geographic mapping methods are becoming more popular (Leydesdorff et al., 2013). A sample of research from leading authors in this area is included in Table 4.

Several researchers have used multi-criteria decision making (MCDM) (Phan, 2013) (Mumpower and Stewart, 1996) (Tran, 2013b) to consider different perspectives in their research. A multi-level decision model (MLDM) is a flexible method that can utilize both structured data and unstructured data by using methods that quantify the expert judgment. Saaty (Saaty, 2008) introduced the Analytic Hierarchy Process (AHP), a popular MCDM method to deconstruct a problem into top-down levels of linked concepts. The Hierarchical Decision Model (HDM) is similar to the hierarchical structure of approaching problem and differs in the use of pair-wise comparisons to quantify element weights.

Hierarchical decision models (HDMs) were developed by Phan to evaluate the innovativeness of companies in the semi-conductor industry based upon output indicators (Phan, 2013) and by Tran to develop an index to measure the effectiveness of a technology transfer office (TTO) based upon fulfillment of the stated organizational mission (Tran, 2013b). These researchers measured a broader range of outcomes to include knowledge transfer beyond licensing. In Tran's research, a knowledge and technology transfer effectiveness index was developed to compare mechanisms for a particular university. This research is particularly interesting for this study because it provides precedence in the literature for using the HDM as an appropriate methodology as well as additional data to identify knowledge and technology transfer output elements.

The multi-attribute utility theory (MAUT) is another popular multi-criteria model that considers additive value for multiple objectives (Iskin, 2014). Because the AHP and the HDM involve a relative importance assessment procedure and use "a hierarchy to establish preferences and orderings" they are "sometimes classified as a MAUT approach" [(Wallenius et al., 2008) p 646]. The MAUT process considers the perspective of a decision maker through the use of utility functions or desirability curves.

Literature clearly documents the importance of CRCs for

Table 3
Case-based research in CRC literature.

Author year	Focus	Findings	Gaps
Gray and Steenhuis (2003)	IUCRC Evaluation process	Centers have been extensively evaluated	Comparative evaluation missing or of low quality
Corley et al. (2006)	Multi-institutional research evaluation implications	Need organizational structure or epistemic development of the disciplines in the collaborations	More focus needed on the design of organizational systems.
Gray (2008)	IUCRC Evaluation Process	Structured case reports needed to include outcomes, best practices and breakthrough technologies	Subjective data are non comparable, coding methods needed
Ramanathan et al., 2013)	CETI IUCRC Stakeholder needs assessment	Agile design processes benefit students to span boundaries	Innovation outcomes are typically unmeasured
Scott, 2014 (Scott, 2014)	IUCRC break- through technologies	IUCRCs need a structured way to report breakthrough technologies	Inconsistency of impact data.

Table 4
SNA research in CRC literature.

Author, year	Topic
Balconi and Laboranti (2006)	University-industry interactions in applied research: The case of microelectronics
Rafols et al. (2009)	Science overlay maps: a new tool for research policy and library management
Porter and Rafols (2009)	Is science becoming more interdisciplinary? Measuring and mapping six research fields over time
Abbasi and Altmann (2011)	Correlation between Research Performance and Social Network Analysis Measures Applied to Research Collaboration Networks
Garner et al. (2012)	Assessing research network and disciplinary engagement changes induced by an NSF program
Leydesdorff et al. (2013)	Global maps of science based on the new Web-of-Science categories
Abbasi et al. (2014)	Measuring social capital through network analysis and its influence on individual performance

Table 5
Example of performance evaluation challenges found in literature.

Reference	Findings	Theme
Boardman and Gray (2010)	"CRCs are inherently complex and therefore a challenging phenomenon to understand". [(Boardman and Gray, 2010) p 5]	Complexity
Roessner et al. (2010)	Lack of a "standardized performance criteria" and "exclusive reliance on quantifiable data" provides misleading results (Roessner et al., 2010).	Traditional indicators inadequate
Schmoch et al. (2010)	"Scientific performance should not be measured by a one-dimensional metric such a publication, since it is a multi-dimensional phenomenon." [(Schmoch et al., 2010) p2]	Traditional indicators inadequate
Palomares-Montero and Garcia-Aracil (2011)	"It is difficult to obtain valid and reliable data and the results of evaluation processes depend on the quality of the information available. There is a lack of disaggregated data to enable comparison among disciplines, and data often are not sufficiently firm, resulting in indicators that provide inaccurate results". [(Palomares-Montero and Garcia-Aracil, 2011) p353]	Lack of available data, Traditional indicators inadequate
Penfield et al. (2014)	"These 'traditional' bibliometrics techniques can be regarded as giving only a partial picture of full impact with no link to causality. (Penfield et al., 2014)	Traditional indicators inadequate
Abbasi et al. (2014)	"Collecting network data has its own limitations" and lack of other types of data prevents performance comparisons. [(Abbasi et al., 2014) p72]	Lack of available data

translational research; but, performance comparison is still somewhat of a challenge (Bruneel et al., 2010; McGowan, 2012). Where formal evaluation programs exist, the methods are typically resource intensive with results focused on a single center or at the program level (Gray, 2008). Table 5 provides evidence for the three leading gaps in the CRC performance evaluation research literature: ecosystem complexity, lack of data, and inadequacy of traditional indicators.

CRC's are complex ecosystems with multiple actors, missions and organizational structures (Schultz, 2012). Basically, "improved methods are needed for program evaluation" [(Ruegg, 2006) p 11] because a GUI CRC is a complex ecosystem (Adner and Kapoor, 2010); not a "trivial machine, with a defined input-output ratio" (Wallner and Menrad, 2004). Additional expert input was obtained through a proposal process for this research. Representatives from the NSF Science of Science & Innovation Policy (SciSIP) program provided additional comments. "*Many federal science agencies support large centers of research around a single scientific problem. These centers can vary considerably in the science they support, their structure, and ultimately their strengths. Where one center may make considerable progress in research, another may instead succeed best at producing excellent scientists. Agencies have long struggled with how to evaluate such centers, given their complexity*."

While traditional outputs of university research projects such as publications and patents are easily quantified with bibliometrics data, "exclusive reliance on quantifiable data" causes misleading results (Roessner et al., 2010) by painting a partial picture (Penfield et al., 2014). However, "identifying a set of metrics to evaluate the performance of a university-based ecosystem was [remains] a considerable challenge" [(Graham, 2013) 4]. Thus, the "STI [science and technology] indicators that were important last century may no longer be so relevant today and indeed may even be positively misleading" [(Freeman and Soete, 2009) p588]. Or worse, are simply the "wrong" metrics (Wallner and Menrad, 2004).

Metrics can be used to compare and differentiate the performance of different organizations. Some organizations produce outputs more efficiently than others or at higher quality levels. Effective use of metrics can help organizations to achieve superior performance outcomes. However, Freeman and Soete argue on the basis of their 40 years of indicators work that "STI [science and technology] indicators that were important last century may no longer be so relevant today and indeed may even be positively misleading" [(Freeman and Soete, 2009) p588]. Researchers have found that a GUI CRC is a complex ecosystem (Adner and Kapoor, 2010); not a "trivial machine, with a defined input-output ratio" (Wallner and Menrad, 2004). So, metrics are important; but, which ones are appropriate?

Publications and patents are common outputs of university research projects. Publications typically represent the output of earlier-stage, basic research while patents are typically more indicative of applied research (Schultz, 2012). These traditional outputs are easily quantified with bibliometrics data and have been used in many studies. However, researchers have cautioned that "exclusive reliance on quantifiable data" provides misleading results (Roessner et al., 2010) because they only provide a partial picture (Penfield et al., 2014). Others have cautioned that traditional measures are simply the "wrong" metrics (Motoyama et al., 2014a; Wallner and Menrad, 2004).

Knowledge transfer and integration also requires understanding of social dynamics and networks. Emerging research in social network analysis and metrics such as betweenness centrality and diversity are promising; but, the use and interpretation is difficult (Wagner et al., 2011). In a recent, empirical research study involving multiple experts the results concluded that "identifying a set of metrics to evaluate the performance of a university-based ecosystem was a considerable challenge" [(Graham, 2013) 4].

Another group discusses challenges attempting to tie the metrics to the outputs and outcomes because more and better quality data are needed to answer impact type of questions (Adams et al., 2001). Some of the available aggregated data was found to be of poor quality leading to inaccurate results [[70]p353]. In general, researchers agree that "due to non-availability of data we are unable to measure" performance of research centers. Researchers are specifically asking for time series membership data (Adams et al., 2001) and network data (Abbasi et al., 2012) that is disaggregated (Palomares-Montero and Garcia-Aracil, 2011).

In summary, performance measurement calls for a comprehensive (Anderson et al., 2007b), multi-dimensional approach considering multiple perspectives. This problem requires boundary-spanning criteria because there are many constituent groups who have a stake placing different values on outputs and outcomes. Different perspectives can lead to disagreement about the mission and value of the outputs. For example, different institutional norms govern public and private knowledge (Geuna and Muscio, 2009) (Bruneel et al., 2010). Even when agreement is reached, stakeholder perspectives are expected to shift over time. Literature is calling for more research to examine the effectiveness of the CRC organization and the impact of their activities and outputs (Carayannis and von Zedtwitz, 2005).

With limited resources, policy makers must be diligent at attempting to make objective and increasingly transparent funding decisions. Despite the importance an increasing investment, a set of holistic output indicators are missing. Missing also are decision support tools and methods to help make performance measurement more cost effective. Without the help of such tools, policy makers are ill equipped to make transparent and objective decisions. They need to know if their program really makes a difference "compared to no program or an alternative one" [(Gray, 2008) p 78] and how to improve with scarce resources. Therefore this paper adds value to the stream of literature by developing a model that measures the degree to which different science and engineering centers meet a program's mission specifications using a balanced set of performance indicators.

3. Methodology

CRC performance should be measured using multi-dimensional criteria because this is a "multi-dimensional phenomenon" [(Schmoch et al., 2010) p2]. Understanding that organizational effectiveness is a construct rather than a concept (Quinn and Rohrbaugh, 2014) helps to explain why a multi-criteria decision making tool is appropriate for this type of a problem. In the organizational theory literature, Steers (Steers, 1975) and other researchers (Cameron, 1978) discuss the importance of using a framework to link decision criteria (Quinn et al., 1981). Concepts are abstractions defined and measured by characteristics. Higher-level abstractions are often difficult to characterize and measure requiring construction of different concepts.

3.1. Multi criteria decision models

Multi-criteria, multi-level models are useful when decisions are complex and require judgment between multiple alternatives. They present an appropriate method for this study for several reasons:

1) They are flexible, decision support tools that can be used to quantify expert judgment. These methods can handle both qualitative and quantitative data.
2) The hierarchical methods allow for decomposition of a complex decision problem into a hierarchy of smaller sub-problems for independent analysis. The elements of the hierarchy can relate to any aspect of the decision problem under investigation.
3) There is a precedence in the literature. The methods have been used in other research studies to explore complex, multi-dimensional problems (Chen and Kocaoglu, 2008a; Mumpower and Stewart, 1996; Phan, 2013; Tran, 2013a).

Cleland and Kocaoglu introduced a mission-objectives-goals-strategies-activities (MOGSA), hierarchical decision model (HDM) framework (Cleland and Kocaoglu, 1981) that is well suited for this performance evaluation problem. A key aspect of this method is that the problem can be broken into a hierarchical structure (Saaty, 2008), where experts can judge a series of elements in pairwise comparisons. Fig. 2 shows how the new model follows the first three levels in the MOGSA framework and replaces the 4th level with measureable outputs.

3.2. Hierarchical decision model

The human brain is designed to analyze complexities by compartmentalizing them and splitting the parts in turn into smaller parts to deal with individually, since it cannot deal with too many factors at the same time. This hierarchical vertical structure is our natural way of thinking. A cross-sectional way of analyzing relations is beneficial when you have a certain objective and want to understand the effect of other factors or the relationship between entities. HDM allows the decision maker to divide the problem into its smaller entities for analysis and therefore reveal any hidden relationship between elements. This methodology has been used for policy planning for a variety of objectives and was proven practical (Hämäläinen and Karjalainen, 1992) (Gerdsri and Kocaoglu, 2008) (Elkarmi and Mustafa, 1993) (Lee et al., 2007) (Lee et al., 2008).

The other advantage of the HDM is the ability to screen and select a large number of alternatives. Also, a large number of criteria and sub criteria can be used, which allows the analyst to cover the topic under investigation from many different angles. The results of the HDM are not just solid numbers or ranking, this model allows the analyst to dig deep into the results and identify other trends or priorities within the same criteria. This will be of great value for the proposed model since policy analysis is not a binary problem, but needs deep analysis of the integrated relationship among objectives, barriers, and benefits.

This approach will be useful to gain insight into current policies and criteria that are constantly changing with the fast pace of technology development, which is not always accounted for in the literature. This research has utilized the HDM methodology which allows for breaking down the problem into a hierarchical structure in order to analyze the relationship between a mission, objectives, and alternatives (see Fig. 2). HDM is used to quantify expert qualitative judgments and convert them to numerical values using a pair-wise comparison method (Table 6).

By using Constant-Sum Method, a total of one hundred points was assigned by experts, divided between any two elements at the same level. For the level of mission (M), quantifying expert judgment relative to the contribution of the objective level to the mission is given as ${C_l}^{O-M}$ (see Table 9 for all model notations). The overall relative contribution of the energy policy alternative (A) to the mission (M) is calculated by adding the sum products of all local contribution matrices between M and A and is given by (Chen and Kocaoglu, 2008b).

$$\boldsymbol{C_i^{A-M}} = \sum_{l=1}^{L} \sum_{k=1}^{K} \boldsymbol{C_l^{O-M}}.\ \boldsymbol{C_{kl}^{G-O}}.\ \boldsymbol{C_{ik}^{A-G}} \tag{1}$$

For each level, the judgments were collected and converted to weights. The alternative with the maximum weight sum would be the best "fit" to the mission. Inconsistency and disagreement metrics (Mumpower and Stewart, 1996; Phan, 2013) were used to ensure robustness of the model.

4. Model development

The purpose of the model (decision objective) is placed at the top of the mission-oriented framework. Organizational objectives fill the second level. Goals are placed at the third level and output indicators used to measure the goals fill in the 4th level. Thus, the mission of the model is a performance evaluation score that determines the degree to which objectives measured by a balanced set of output indicators contributes to the IUCRC program's mission (Fig. 3).

It makes sense that different outputs are not valued the same. Some may contribute to performance more or less than others. The value of relative outputs towards the mission is determined by experts. Mean scores of experts in each panel are then quantified to develop weights for each element. It also makes sense that producing different output quantities meeting different quality standards will provide different results.

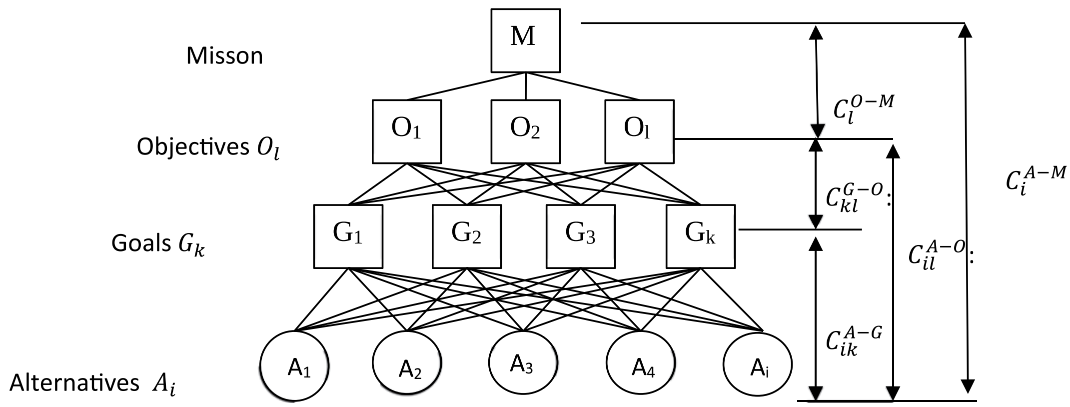


Fig. 2. Generic form of HDM with four decision levels.
(Adopted from 87, 89.)

Metrics developed for each output are valued using desirability curves. More is not always better and scales are not absolute. Thus, curves reflecting desired output quantity and/or quality are developed.

Eq. (2) shows how a performance effectiveness value (E) can be calculated using multiple criteria (c) for any number of (I) alternatives (a) under comparison.

Performance effectiveness value

$$\mathrm{E}\,(\mathrm{a_i}) = \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{j=1}^{J} O_l G_k^l C_{jk}^k\, d\,(m, jk) \text{ for } i = 1, \ \ldots, \mathrm{I} \tag{2}$$

where:

E $(\mathrm{a_i})$ = Effectiveness value for alternative I,
O_l = The degree to which objective $\boldsymbol{l}$ contributes towards center performance.
${G_k}^l$ = The relative contribution of goal $\boldsymbol{k}$ under objective $\boldsymbol{l}$ towards performance.
I = Number of alternatives under comparative evaluation,
J = Number of outputs,
K = Number of goals,
L = Number of objectives,
d $(m_i,\ jk)$ = Metric desired value of alternative (i) for jth criterion under goal (k),
${C_{jk}}^k$ = Relative importance of criterion (jk) under goal (k).

4.1. Expert panels

Expert judgment is a key component in this research approach. Experts validated the linked model elements for content and construct. Expert panels were formed to collect data. After completing data analysis for consistency and disagreement, the accepted data was used to quantify decision element weights finalizing the model. Consultation with experts validated the results and the generalizability of the model. Appendix 1 shows the details on the experts used.

This study uses a two-phased research design where thirty-seven selected experts were formed into five (5) different panels to validate then quantify decision elements. Several experts met the criteria for multiple panels and were motivated to participate on them. Experts in the sixth panel were asked to validate and quantify desired metrics.

Expert numbers were assigned in order that consent forms were received. Many of the experts have multiple titles. The title column is not a complete representation of an expert's experience as many experts fill multiple roles. The primary background qualifying the expert for the study was classified as a regular or contracted employee of the NSF (NSF), a leading researcher (R), or a center director, co-director or executive (C).

Each panel was configured to consider a balanced perspective to minimize bias and encourage a richer and more diverse pool of data Column 1 in Fig. 4 shows how the thirty-seven (37) experts were configured into six (6) panels. Columns 2 and 3 discuss how experts were asked to validate and quantify different levels of decision criteria.

For example, experts in panel 2 validated and quantified goals relative to each of the three (3) objectives. Qualifications for each of the panels and the data collection methods used are also discussed. Separate judgment quantification instruments were created for each of the functions: validation, quantification and desirability curve development. The expert panel formation process also considered how different perspectives are required at three (3) different levels.

Table 7 shows how expert judgment is an appropriate method to validate the model content, construct and results [(Tran, 2013) p71].

4.2. Model components

The hierarchical decision model (HDM) provides a flexible, hierarchical structure for decision analysis. The purpose of the model is to determine the degree to which an IUCRC meets the program's mission. It is a generalizable model that outputs a performance evaluation score for an IUCRC in the program by evaluating a holistic set of metrics.

At the top of the model, the objective is the organizational effectiveness score. At level 2, the NSF IUCRC program objectives specify the

Table 6
Notations for HDM.

Where:		
O_l: Objectives, l = 1,2, …,l	${C_l}^{O-M}$: relative contribution of the Lth objective to the mission	
G_k: Goals, k = 1,2, …,k	${C_{kl}}^{G-O}$: relative contribution of the kth goal to the Lth objective	
A_i: Alternatives, i = 1,2, …i	${C_i}^{A-M}$: Overall contribution of the ith alternative to the mission	${C_{ik}}^{A-G}$: relative contribution of the ith alternatives to the kth goal
		${C_{il}}^{A-O}$: relative contribution of the ith alternative to the kth objective

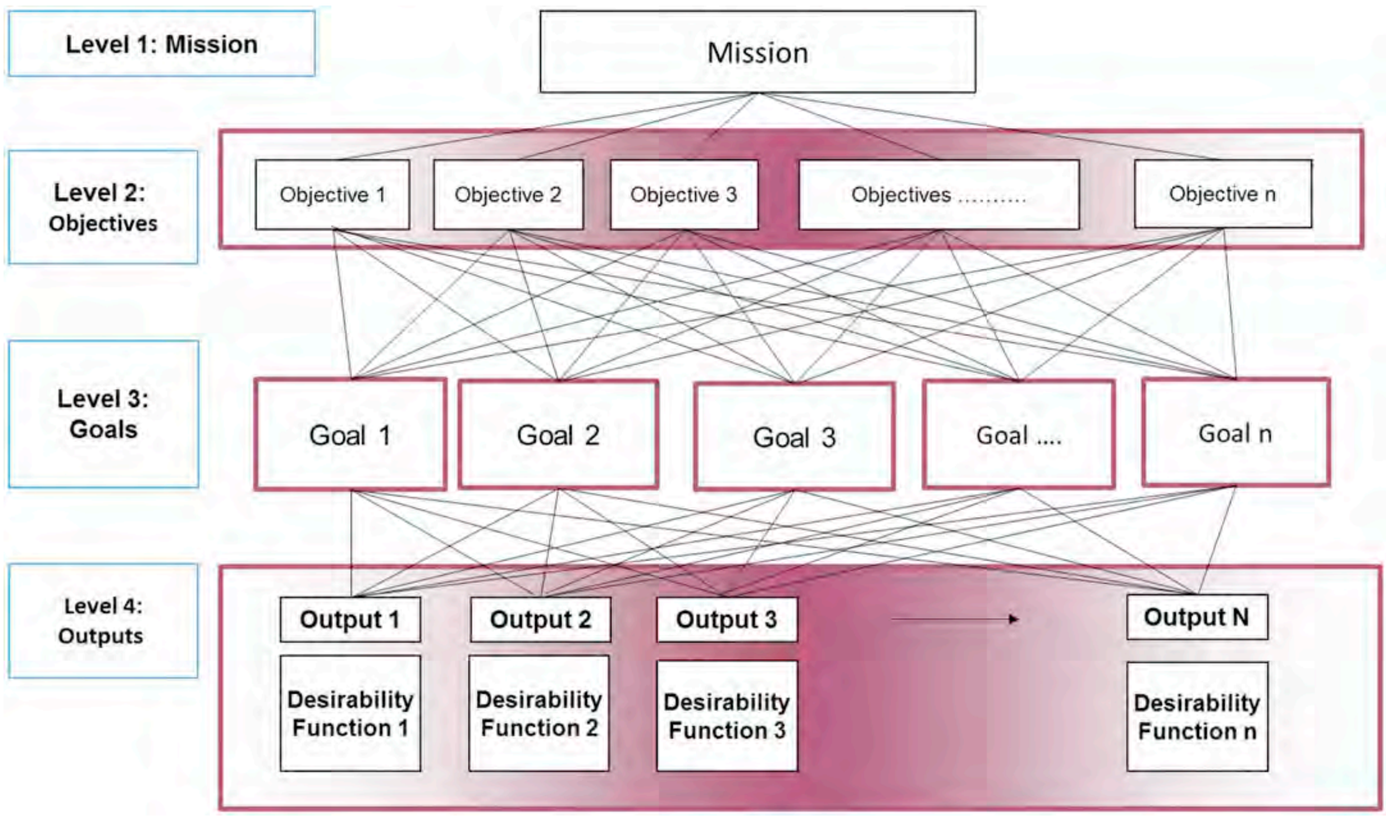


Fig. 3. Generalized hierarchical framework.

mission of the program.

Literature finds the NSF IUCRC program's mission, outlined in what has become known as "the purple book" (Gray and Walters, 1998), is specified by three objectives:

1. To pursue fundamental (collaborative and pre-competitive) engineering and scientific research having industrial relevance.
2. To produce graduates who have a broad, industrially oriented perspective in their research and practice.
3. To accelerate and promote the transfer of knowledge and technology between university and industry (public) ([39]p 23).

The objectives are placed at the second level of the model as shown in Fig. 5.

ID #	Phase 1 (P1): Validate content and construct.	Phase 2 (P2): Data collected to quantify elements.	Expert Panel Qualifications	Research Instruments and Data Collection Methods
1 P1(2) P2(1)	Validation of construct and content of level 2: political.	Quantification of objectives at the political level of the model.	NSF IUCRC directors (past/present) NSF IUCRC Project Director	Phone Interview: open- ended and closed-ended questions for quantification and qualitative contextual data later used for scenario analysis
2 P1(9) P2(8)	Validation of construct and content at the organizational level.	Quantification of goals at the organizational level of the model.	NSF IUCRC program evaluations. Leading researchers in the field.	Four instruments were developed for each panel group for phase 1 (P1) to validate elements respective to each groups perspective. Another set of instruments were developed for phase 2 (P2) to extract pair-wise comparison data for quantification. P1) Binary checklist (agree/ disagree) with open text box for additional qualitative data. P2) Survey requesting pair-wise comparisons with drop-down box allowing for numerical selection of values [1,99].
3 P1(8) P2(14)	Validate sub-criteria outputs relative to parent goals.	Quantify outputs relative to goals characterizing objective 1 (research).	NSF evaluation members, Leading researchers	
4 P1(11) P2(13)	Validate sub-criteria outputs relative to parent goals.	Quantify outputs relative to goals characterizing objective 2 (students).	NSF IUCRC Center directors (past/present), Leading researchers	
5 P1(9) P2(15)	Validate sub-criteria outputs relative to parent goals.	Quantify outputs relative to goals characterizing objective 3 (KTT).	Leading researchers, Center directors/co-directors, IAB Members, KTT experts	
6 P1(4) P2(4)	Validate metrics and data sources available for desirability curves.	Quantify "goodness" data to develop desirability curves.	NSF Evaluators, leading authors respective to perspective: coop. research, IUCRC student, or KTT.	Phone Interviews: Acceptance of metric and context, subjective data to qualify statements for quantification of desirability values. Email.

Fig. 4. Panel configurations.

Table 7
Summary of evaluation tests.

Validity	What is measured	Methods
Construct	The degree to which a measure relates to expectations formed from theory for hypothetical construct	Judgmental, Correlation, Convergent-discrimination Factor analysis Multitrait-multimethod
Content	Degree to which the content of the items adequately represents the universe of all relevant items under study	Judgmental
Criterion-related	Degree to which the criterion can capture the true value of the variable	Judgmental, Correlation

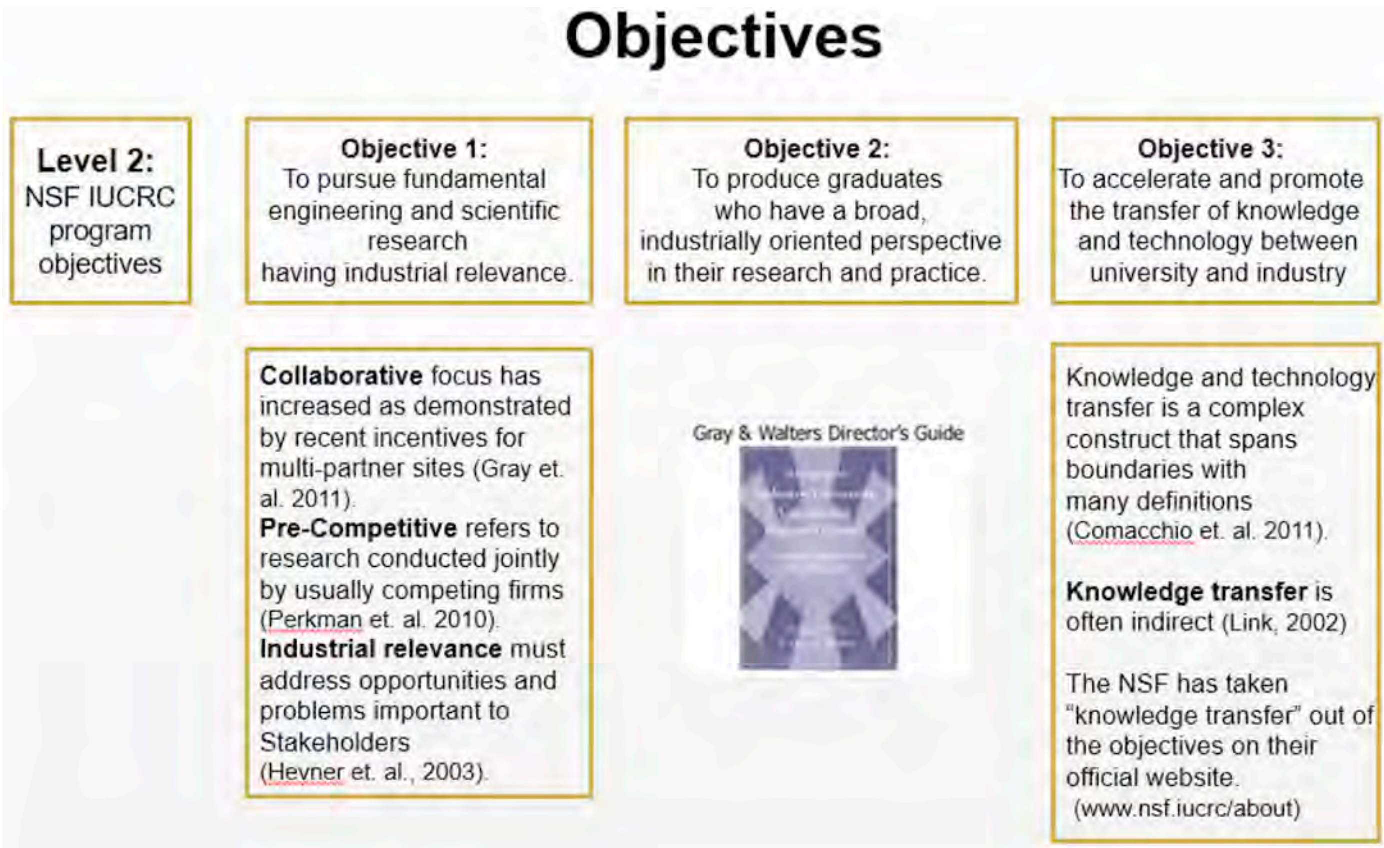


Fig. 5. NSF IUCRC program objectives.

The development of collaborative, pre-competitive research has been a part of the program's mission since inception (Gray and Walters, 1998). Thus, key to the program is promoting boundary spanning activities through cooperative partnerships and multi-disciplinary science (Sundstrom & Gray, 2010). Since the early 1990's, the IUCRC solicitations have increased incentives for multi-site IUCRCs (Gray et al., 2011). The minimum threshold for a multi-site proposer is $350 K while single-site membership requires $400 K per year. A program expert confirmed that a lower threshold for multi-site membership agreements will likely continue.

An IUCRC requires graduate student involvement (Gray and Walters, 1998). Funding and scholarships provide graduate students opportunities to complete research towards a thesis or dissertation making programs more attractive (Behrens and Gray, 2001). Students gain experience and acquire knowledge through a cooperative and industry-oriented approach to conducting research.

Knowledge and technology transfer (KTT) is a complex construct, spanning boundaries (Comacchio et al., 2011) with many definitions. The facilitation of knowledge and technology transfer (Etzkowitz and Leydesdorff, 2000a) is key to achieving the NSF IUCRC's mission as stated in the third objective: to accelerate and promote the transfer of knowledge and technology between university and industry [(Gray and Walters, 1998) p23] that benefits the public (Devine et al., 1987; Feldman and Kelley, 2006b; Roessner et al., 2010).

Each of the three objectives are further characterized by two measurable goals. "New knowledge" and "stakeholder satisfaction" measure how fundamental research is pursued and how satisfied stakeholders are with this pursuit. Producing graduates requires involved students and a strong development program. The goals used to characterize KTT are based on Bozeman's "Contingent Effectiveness Model of Technology Transfer" [(Bozeman, 2000) p 637]. Fig. 6 shows how two goals are linked to each of the three objectives and how the Bozeman model is adapted for this research.

It is important to carefully select outputs (Piva and Rossi-Lamastra, 2013) that not only "fit" the mission specifications but are also aligned with the social technology characterizing the NSF IUCRC program. Experts provided qualitative input regarding the ability of decision elements obtained from the literature review to represent the uniqueness of the NSF IUCRC program. Then, experts judged each element providing quantitative binary acceptance data using a Delphi process.

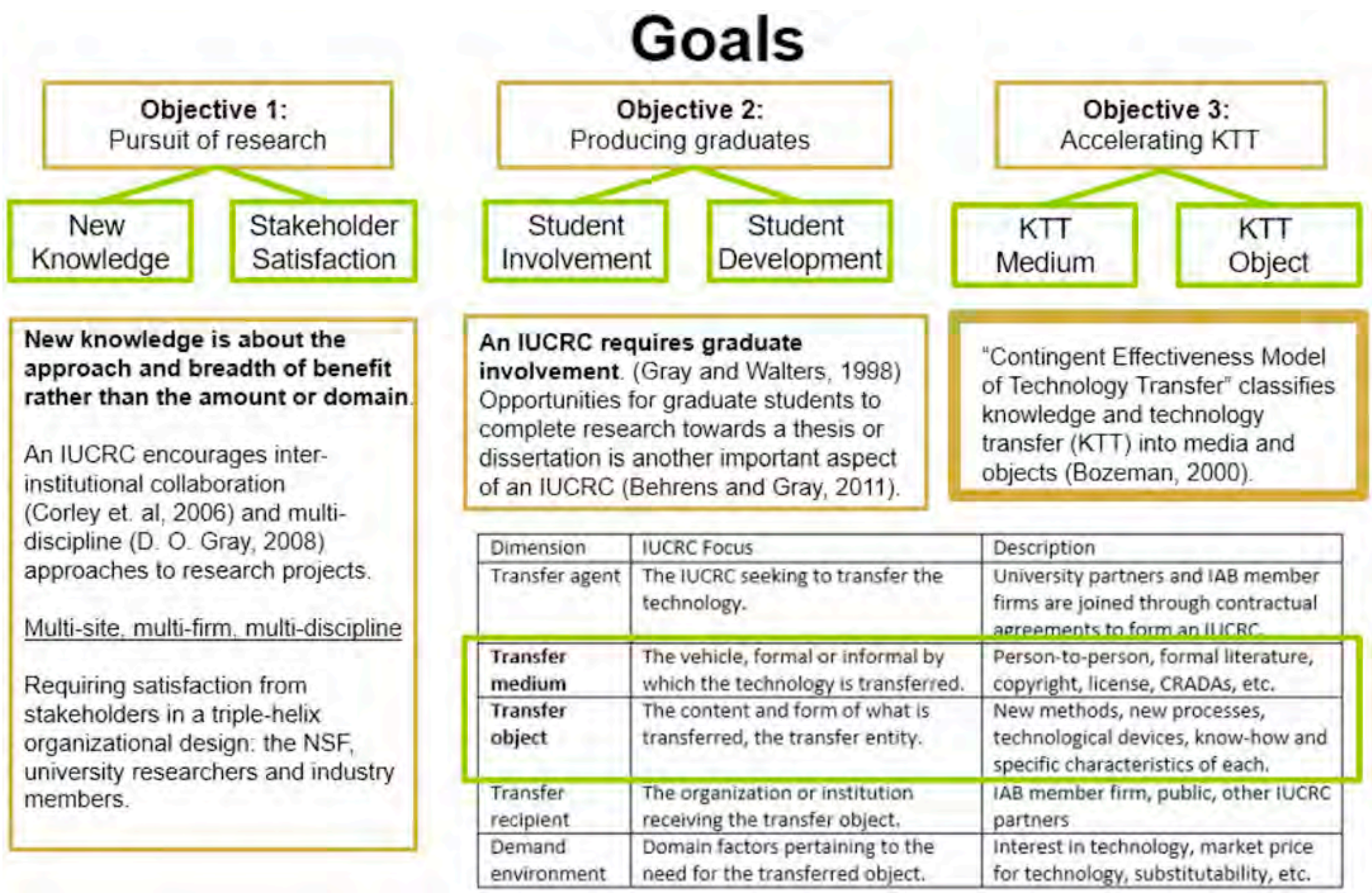


Fig. 6. Goals.

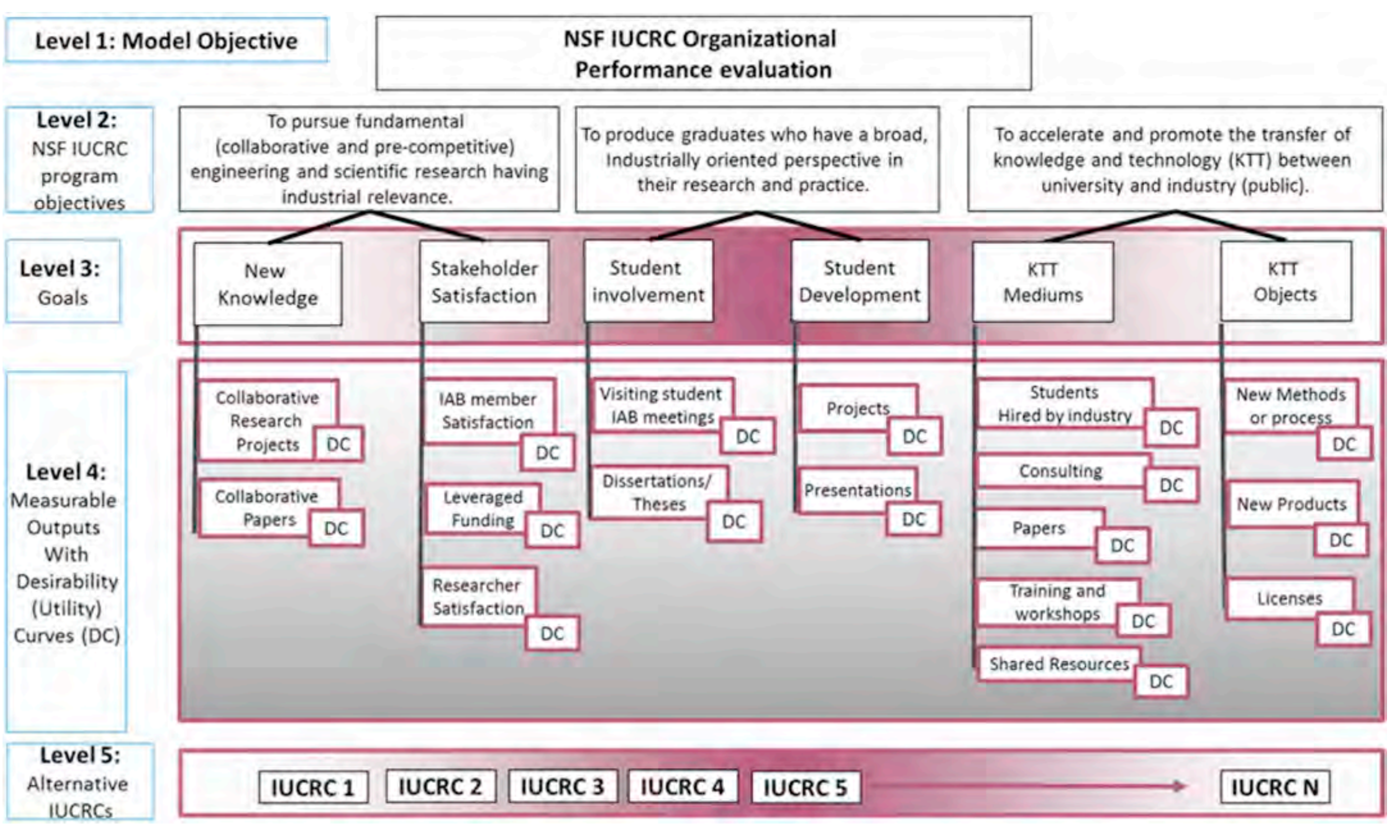


Fig. 7. Validated model construct.

Table 8
Literature identified new knowledge outputs.

New knowledge	IUCRC focused description	Reference
Scientific co-publications	Co-authorship. The IUCRC literature emphasizes authors to be affiliated with different organizations.	(Gibson, 2015) (Gray et al., 2012b) (Rohrbeck and Kaab, 2013) (Bozeman, 2000) (Gray et al., 2013) (Gray, 2008) (Chukumba and Jensen, 2005; Kostoff, 2005) (Shane and Somaya, 2007b)
Patents and co-patenting	Multiple firms listed as owners.	(Grupp and Linstone, 1999) (Cunningham et al., 2014) (Rohrbeck and Kaab, 2013) (Bozeman, 2000) (Gray, 2008) (Kostoff, 2005) (Cohen et al., 1994)
Collaborative research projects	Researchers affiliated with multiple organizations. Multi-disciplinary research has been recently emphasized in the IUCRC and team science literature.	(Cunningham et al., 2014; Etzkowitz and Leydesdorff, 2000a) (Gray et al., 2012b) (Rohrbeck and Kaab, 2013) (Bozeman, 2000) (Gray, 2008) (Kostoff, 2005) (Cragin et al., 2012) (Shane and Somaya, 2007a) (Shane and Somaya, 2007b) (Hicks, 2012) (NSF, 2013)

Elements were accepted when an 80% agreement level was reached by the panel of experts (Tran, 2013a). The validated model is shown in Fig. 7 and used to guide this discussion.

Literature documented outputs for new knowledge generated through multi-disciplinary and multi-site collaboration are summarized in Table 8. Through the validation process, experts clarified that "*patents are explicitly NOT a part of the IUCRC program*" because they discourage pre-competitive research.

In an IUCRC there are three primary stakeholder groups: government, university and industry. The satisfaction of each group must be considered; however, this is somewhat of a challenge because often there are competing needs. For example, researchers seeking tenure may be motivated to publish and become frustrated if an IAB member lobbies for publication to be delayed. Some IAB members may be short sighted and not appreciate the nature of pre-competitive research, instead being more focused on solving an immediate problem facing their company. Industrial advisory board (IAB) members can be satisfied in an IUCRC that is not performing well if they are getting more benefit from the research. So, it is important to consider the trade-offs among the three primary stakeholder groups.

There was some debate about student involvement, participation and engagement at IAB meetings. Some IAB meetings have allowed members to attend using electronic communications. With advances in today's communication technologies such as video conferencing, some consider remote attendance at meetings as sufficient. However, researchers have found the value of long-distance participation to be limited (Sundstrom & Gray, 2010).

While literature identifies many different outputs for student development such as: number of courses taken, number of degrees earned, number of projects completed, papers written and presentations given; the IUCRC program is focused on research and presentations. Students will receive degrees whether they conduct industry-related research or not.

Bozeman describes a KTT medium as the vehicle, formal or informal by which the technology is transferred (Bozeman, 2000). KTT media supported by literature include personnel exchanges, demonstrations, papers and professional networks. Shared knowledge and idea generation (Gray et al., 2011) transferred at networking and informal events are difficult to evaluate often using attendance and participation as proxy measures. Knowledge generation and knowledge transfer is evaluated differently. When students, faculty or industry members conduct research they are creating knowledge whereas when they are teaching or taking a course they become containers for knowledge. Table 9 summarizes KTT media sources defined in the literature.

KTT objects provide the form and document the content of what is transferred. Some examples of this transfer entity include new products or services, new methods or processes and patents. In an IUCRC, focus is placed on a technological breakthrough or advance such as: "significant process improvements, new process or techniques, and new or improved products or services that resulted either directly from, or was indirectly stimulated by the center's research program" (Scott, 2014). The NSF has published a set of Compendiums that catalogue peer reported breakthrough technologies. Table 10 identifies KTT objects found in literature.

Table 11 shows how metrics are used to describe each output. The parent element for each output is a relative goal that is identified in column 1.

Development of a desirability curve is a method to convert either qualitative or quantitative data used for measuring a decision element to a scaled quantitative value. Understanding the desired or ideal value for a metric is important. The relationship of values for different metrics may scale differently. Comparing desired values against a consistent scale normalizes the values.

So, what value is desirable for each of the outputs? In a complex ecosystem, stakeholders may provide conflicting judgment about these values. For example, IAB membership renewal rates are used to measure IAB member satisfaction. If experts agree that some turn-over is normal and a desired retention is 80% or better, 40–50% retention may or may not be judged to be half as good. A 60% retention may signify a tipping point or problem.

4.3. Final model

Fig. 8 shows the model quantified through the expert panels with HDM. This model was applied to a case study in the next section.

5. Case study application

A case study is developed to illustrate how the model works and to conduct criterion-related validation. Criterion-related validation enlists the help of an expert to evaluate the degree to which the model reflects actual performance. Data collected for the Wood Based Composites (WBC) center was used to populate the metrics, find respective desirability values and calculate a score. Consultation with experts validated the results and generalizability of the model.

5.1. Case study background

The mission of the Wood-Based Composites Center (WBC) (Fig. 9) is to advance the science and technology of wood-based composite materials. While the center was formed with only 2 partner universities, it has grown to informally include four more. On their website (wbc.vt.edu) the center discusses goals that include attracting students to careers in the wood-based composites and adhesion industries by providing "intellectual exchange and interaction among professionals and students."

Data was collected from five secondary data sources: center websites, NSF IUCRC structural information reports, center minutes, the NSF Compendium of Breakthrough Technologies and the ProQuest and interviews. Information about collaborative projects and background information on researchers and configurations of projects was obtained from specific IUCRC websites. From the NSF IUCRC evaluation program database, structural information reports from 2010 to 2014 were used for most of the descriptive statistics. Data regarding attendance was

Table 9
Literature identified KTT media.

KTT media	IUCRC focused description	References
Papers	Publications in peer-reviewed journals are traditionally recognized outputs of KTT.	(Gibson, 2015) (Gray et al., 2012b) (Rohrbeck and Kaab, 2013) (Bozeman, 2000) (Gray et al., 2013) (Gray, 2008) (Chukumba and Jensen, 2005; Kostoff, 2005) (Shane and Somaya, 2007b) (Gaughan and Corley, 2010) (Boardman and Corley, 2008; Geisler, 2010; Ponomariov and Boardman, 2010)
Reports	Research reports	(Kostoff, 2005) (Shane and Somaya, 2007b)
Conference presentations		(Etzkowitz and Leydesdorff, 2000a) (Gray et al., 2012b) (Rohrbeck and Kaab, 2013) (Gray et al., 2013) (Gray, 2008) (Kostoff, 2005) (Cragin et al., 2012) (Hicks, 2012; Shane and Somaya, 2007b) (Cohen et al., 1994)
Workshops, classes	Attendance at IUCRC directors meetings and IAB meetings, workshops.	(Boardman and Bozeman, 2015) (Etzkowitz and Leydesdorff, 2000a) (Kostoff, 2005) (Cragin et al., 2012) (Bruneel et al., 2010) (Hicks, 2012) (Cohen et al., 2002) (Gaughan and Corley, 2010; Perkmann et al., 2013) (Behrens and Gray, 2001)
Informal meetings	Informal meetings, one-on-one discussions or small informal groups	(Gray et al., 2012b) (Bozeman, 2000) (Gray et al., 2013) (Kostoff, 2005)
Professional networks: Editors, Professional Organization officers, Boards	Editorships and members in scientific advisory boards and officers of professional organizations improve linkages and the profile of the organization. Editors often find knowledgeable referees who agree to review papers, officers organize conferences and meetings.	(Carayannis et al., 2014b) (Gray, 2008) (Shane and Somaya, 2007b) (Hicks, 2012) (Gray et al., 2012a; Schmoch et al., 2010) (Etzkowitz, 1998; Gaughan and Corley, 2010; Rivers, 2010) (Gray et al., 2001; Tsai, 2015)
Graduate hires, fellowships	Graduates hired into the industry	(Etzkowitz and Leydesdorff, 2000a) (Gray et al., 2012b) (Rohrbeck and Kaab, 2013) (Bozeman, 2000) (Gray, 2008) (Kostoff, 2005) (Chukumba and Jensen, 2005) (Shane and Somaya, 2007a) (Hicks, 2012) (Hayton et al., 2010) (NSF, 2013)
Co-supervising	Supervisors from multiple sites or multiple organizations	(Gray et al., 2012b) (Rohrbeck and Kaab, 2013) (Gray, 2008) (Cragin et al., 2012)
Personnel exchange	Focus on student internships, mentorships.	(Gray et al., 2012a) (Bozeman, 2000; Gray, 2008) (Kostoff, 2005) (Shane and Somaya, 2007b) (Hicks, 2012) (Cohen et al., 1994) (Bozeman, 2000; Cohen et al., 2002; Hicks, 2012; Rogers and Bozeman, 1997; Tyler, 2013) (Bozeman and Boardman, 2013; Perkmann and Walsh, 2007; Tran, 2013a)
Consulting services	Secondary focus on scientific faculty contracted by IAB member firm to facilitate commercialization of technology.	(Etzkowitz and Leydesdorff, 2000a) (Cunningham et al., 2014) (Gray et al., 2012b) (Bozeman, 2000) (Gray et al., 2013) (Gray, 2008) (Kostoff, 2005) (Cragin et al., 2012) (Etzkowitz and Leydesdorff, 2000b) (Hicks, 2012) (Cohen et al., 1994)
Shared resources	Examines not only alternative uses of resources but also possible impacts on the mission such as improved human capital for conducting future research	(Gray et al., 2012b) (Gray, 2008) (Kostoff, 2005) (Hicks, 2012) (Sundstrom & Gray, 2010; Boardman and Corley, 2008; Chai and Shih, 2016; Katz and Martin, 1997)

collected from the NSF evaluator. The Compendium of Breakthrough Technologies provided data regarding new methods and processes. The ProQuest database was searched to identify theses and dissertations published by students with advisors affiliated with IUCRC research projects. A content analysis was conducted on the abstract and the acknowledgement section of each identified student thesis or dissertation to ascertain if the research topic was aligned with an IAB research topic.

5.2. Data collection

The next step is to populate each metric with the data. A metric (m) for an output criteria (c_j) under the jth goal with respect to the kth objective can be represented as (m_{WBC}, jk). The metric for collaborative papers is used to illustrate how the data from the NSF database can be collected to obtain an actual value. Eq. (2) uses data collected from the last three available NSF Structural Information (SI) reports to calculate the number of renewed IAB memberships.

IAB member renewal

$$IAB\ member\ renewal = (\#members\ starting - \#of\ members\ left) \tag{2}$$

Eq. (3) uses this formula to calculate a metric value for IAB member satisfaction using the percent of members who renew.

Percent member renewal

$$\%member\ renewal = (\#IAB\ member\ renew)/(\#starting)$$

$$(m,\%member\ renewal) = \frac{\left[\left(\frac{8}{8}\right)+\left(\frac{8}{9}\right)+\left(\frac{9}{11}\right)\right]}{3} * 100 = 90.2\% \tag{3}$$

The results of the data collection for each metric, (m, jk), are presented in Table 12. The metric and its relative jth criterion are identified in the first two columns followed by the resulting value obtained

Table 10
Literature identified KTT objects.

KTT objects	IUCRC focused description	References
Licenses	Traditional indicators long used in the literature to measure technology transfer. Often an indicator of intent to commercialize the technology.	(Grupp and Linstone, 1999) (Bozeman, 2000; Cunningham et al., 2014; Rohrbeck and Kaab, 2013) (Gray, 2008) (Kostoff, 2005) (Etzkowitz and Leydesdorff, 2000b; Shane and Somaya, 2007a; Shane and Somaya, 2007b) (Cohen et al., 1994) (Feldman and Kelley, 2006a) (NSF, 2013)
New products New methods or processes	Focus on pre-competitive and collaborative. Beneficial to industry (beyond 1 company) Compendium of breakthrough technologies compiles a list of new products and methods by IUCRC (Roessner, 2000).	(Etzkowitz and Leydesdorff, 2000a; Gray et al., 2012a; Gray et al., 2014c; McGowan, 2012)

Table 11
Output decision elements with metrics.

Goal	Output	Description	Metric
New knowledge	Collaborative research projects	An IUCRC encourages inter-institutional collaboration (Corley, Boardman, & Bozeman, 2006) and multi-discipline (D.O. Gray, 2008) approaches to research projects.	% collaborative research projects/3 yr
	Collaborative papers	Pre-competitive research is reflected when IAB members co-author papers.	# collaborative papers/3 yr
Stakeholder satisfaction	IAB member Satisfaction	Gray found membership renewal to be a good proxy for relevant research (D. Gray, Lindblad, & Rudolph, 2001).	% IAB member renewal/3 yr
	Leveraged Funding	NSF requires a minimum membership requirement. Increased leverage over time is desired.	$ total/$ NSF
	Researcher Satisfaction	Coberly found retention, involvement and amount of funding to be highly correlated indicators of researcher satisfaction (Coberly, 2004).	% researcher retention/3 yr
Student involvement	Visiting students IB meetings	Partner site student in-person attendance at IAB meetings. Electronic forms of “attendance” diffuse the social technology and reduce student benefits.	% off-site student attendance at meetings/3 yr
	Dissertation/thesis topic	# of student dissertations or theses using IAB project as research topic.	# of students using IAB research as topic
Student development	Projects	Students are a requirement of the program. Industry acknowledges that access to a diverse pool (Porter and Rafols, 2009) of experience students trained in industrial relevant research is desired and beneficial (McGowan, 2012) (Elena-Pérez, Saritas, Pook, & Warden, 2011).	# students/3 yr
	Presentations	Presentations at IAB meetings or multi-purpose poster sessions are ways students can practice and develop communication skills. Often, students present to the IAB developing skill while gaining insight from collaborators (Boardman and Bozeman, 2015), (Behrens and Gray, 2001).	% student presentations
KTT mediums	Students hired by industry	When students are hired into industry, knowledge transfers from the higher education institution to the company (McGowan, 2012).	# Univ grads hired by industry/3 yr
	Consulting (contract or exchange)	Academic engagement of scientific faculty contracted to help commercialize research outputs through consulting (Perkmann et al., 2013) is important.	% of scientific staff consulting to industry/3 yr
	Papers	The ratio of number of papers published to the number of researchers in an IAB.	# published/researcher
	Training and workshops	Workshops (Kerry et al., 2013) and informal meetings (Cohen et al., 2002) have been found by other researchers to provide the greatest form of knowledge transfer.	% attendance*# held/3 yr
	Shared resources	Sharing of resources (Corley et al., 2006) such as labs (Teshima and Rung, 2010), testing facilities (Sundstrom and Gray, 2010) and open source software improve “sticky and difficult” (Mowery, 2003)	#labs, #equipment availability
KTT objects	New methods	Focus on pre-competitive and collaborative. Beneficial to industry (beyond 1company)	#
	New products	Compendium of breakthrough technologies compiles a list of new products and methods by IUCRC [187].	#
	Licenses	Traditional indicators long used in the literature to measure technology transfer. Often an indicator of intent to commercialize the technology.	#

from the listed data source.

The value of each metric (m,jk) is standardized using a desirability function. The illustration for the percent of IAB member renewal is continued to show how a desirability curve can be used to standardize a value d(m, jk), for each respective decision criteria. Fig. 10 shows how the calculated value of a 90% renewal rate is very close to a value 100% desired by the experts. In fact, it is closer to 100% than if every member had renewed. Experts expect some turn-over because some smaller companies are sponsored by the SBIR program. While it may be concerning when larger long-term IAB members do not renew, turn-over of smaller SBIR sponsored organizations is desired. Eq. (10) shows how the desired value for WBC's membership satisfaction rate ($c3$) relative to the goal of stakeholder satisfaction ($g2$) is $d(m_{WBC}, c_{3,2}) = 0.97$. Appendix 2 lists all the desirability curves.

Desirability value for membership renewal.

$$\frac{92-100}{100-85}=\frac{x}{90}, \qquad x=-4.8 \qquad 92\%-(-4.8\%)=96.8\% \tag{4}$$

5.3. Results

Populating the rest of the metrics with data yields the desirability values recorded in Table 13.

A final score can be calculated by summing the product of the values found for each $d(m,jk)$ and the decision element's ($C_{jk}{}^{k}$) weight (w_j). Eq. (5) shows the expression used to calculate the sum the products of the two vectors.

Performance evaluation score

$$\sum_{j=1}^{17}[w_j * d(m,jk)] \tag{5}$$

Table 14 reflects the results of applying the expression identified in Eq. (5).

Outputs contributing most to this center's performance include research that translates into new methods, engaged students presenting on research and satisfied NSF and IAB member stakeholders. Areas identified for improvement include the number of graduates selecting IAB research topics for their PhD dissertations or Master's theses and more collaboratively configured research project teams.

The data for this center shows there were no theses or dissertations published by students using topics from the IAB center during the last 3 years of data. The desired value for no publications is 0.03 versus a score of 0.42 for a center with an average of 1 publication/year. The result of encouraging students to use center topics for their PhD dissertation research or Master's Thesis would reflect 5% increase in total performance contribution. On the other hand, increased emphasis, expenditure in time and resources on improving licensing would only improve the score by 1%.

An example of how a reasonable set of actions could impact the overall performance of the WBC to the IUCRC program's mission is provided in Table 15. Note how encouraging students to select IAB topics for their dissertation or thesis could gain the center a 5% increase in overall performance.

As shown, a strength of the model is that the more important decision criteria can be identified and their impact can be analyzed relatively quickly. This can be a powerful aid to managers and policy makers because transparency can lead to better decisions. The model

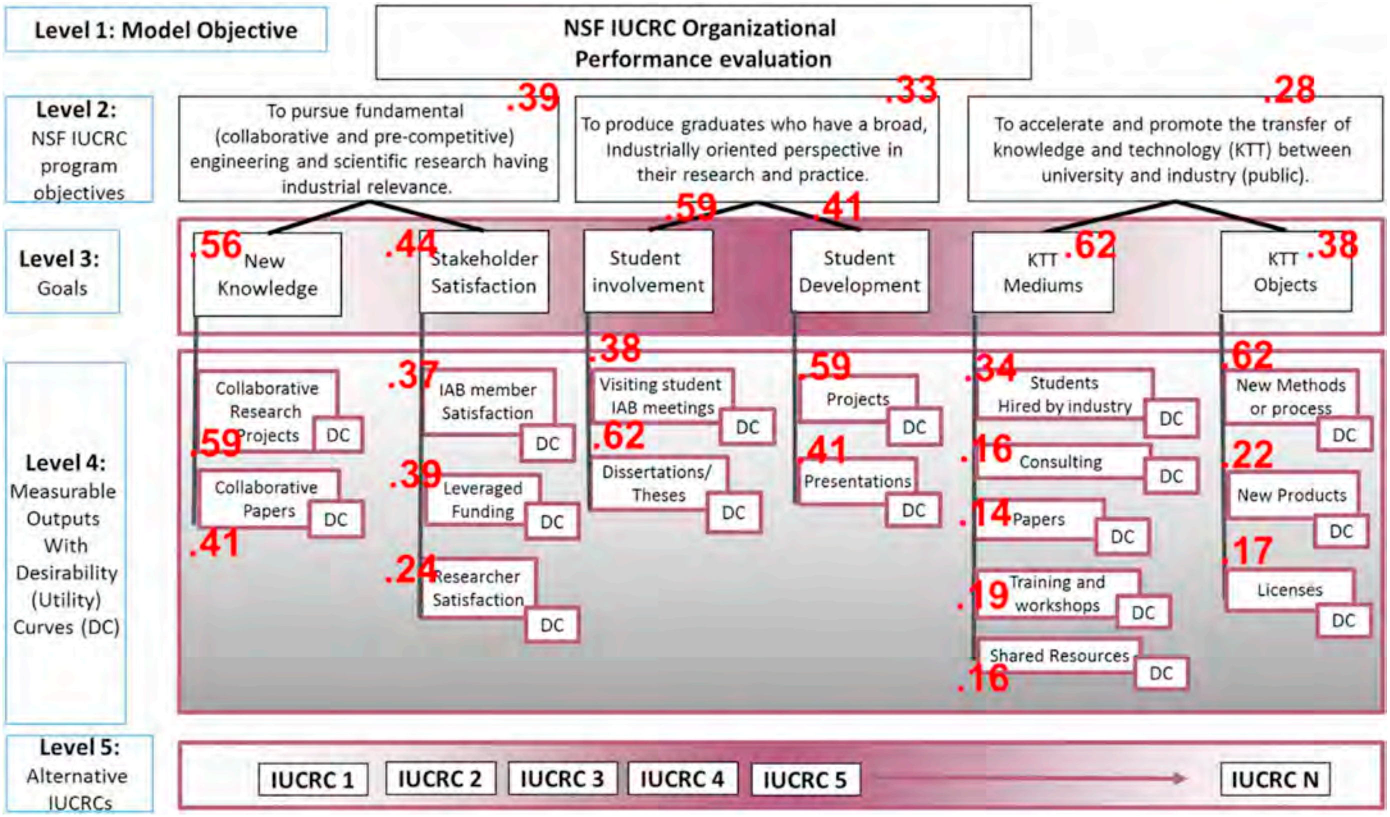


Fig. 8. Generalized HDM for IUCRC performance evaluation.

was sued for another 5 centers and recommendations for them are listed in Appendix 3.

First, experts validated the model's content and construct through a structured Delphi process. Next, expert review of the case study results determined that the model is appropriate and generalizable. Table 16 summarizes how the research design used expert judgment to evaluate results for content validity, construct validity and criterion-related validity.

Experts validated the decision criteria and relative linkages for seventeen elements when an 80% agreement level was met (Tran, 2013a). At level 2 in the HDM, the first objective was changed to emphasize that fundamental research is collaborative and pre-competitive. While the objectives were accepted with these minor changes, experts revealed a healthy level on-going debate about the third objective, knowledge and

Fig. 9. Wood-based composites center.

Table 12
Metric values for WBC test case.

j	Metric	Value	Data source	Approach used
1	% collaborative projects	0.33	Center website wbc.vt.edu	Current number of collaborative project configurations/Total number of projects listed
2	# of collaborative papers	0	NSF www.ncsu.edu/iucrc/	Average number of collaborative papers published as recorded/3 years
3	% IAB member renewal	0.90	NSF	Calculated 3 year average using (members renewed)/members starting
4	Leverage funding ratio	3.83	NSF	Calculated 3 year average using total funding/NSF IUCRC funding
5	% research faculty (RF) change	1.11	NSF	3 year average change for Current number RF/past year number RF
6	% student meeting attendance	0.33	NSF IUCRC evaluator	Averaged for 2 IAB meetings (# non-site students/# total non-site students)
7	% students topics	0	ProQuest database	3 year average (# dissertations or theses published/# students)
8	Student supervision ratio	1.2	NSF	Calculated 3 year average students/RF
9	% students presented	0.14	NSF IUCRC evaluator	# students who presented/# students
10	# students hired	2	NSF	3 year average students hired
11	% RF contracts	0.07	NSF	3 year average RF contracts using in-kind personnel support
12	# Papers published	0.63	NSF	3 year average papers published/researcher
13	% RF meeting attendance	8.87	NSF IUCRC evaluator	2 mtg. average: # RF attending IAB meeting/# total RF
14	Shared resources available	Both	NSF	Binary "yes/no" availability of facilities or equipment
15	# New methods or processes	1	NSF compendium	# reported in recent past compendium
16	# new products	0	NSF compendium	# reported in recent past compendium
17	# new licenses	0	NSF evaluator	Calculated proxy: dependent value based upon new products

technology transfer. This objective has been narrowed on the NSF's website lending more emphasis towards direct commercialization by removing the word "knowledge." However, this focus shift is not supported in the current literature stream or by the experts in this study. Rather, literature emphasizes the importance of knowledge and technology transfer because indirect transfer is often overlooked (Siegel et al., 2004).

In discussing the weighted values of the output decision elements, one expert shared they have "*been concerned for some time about the over emphasis of using licensing and papers as indicators*." Specifically, several experts stressed that "*knowledge and technology transfer is not about the short-term gain of licenses or products developed by one firm, it's really about the long term impact of students who make their career in the field.*" This research supports this viewpoint. For example, a large amount of time and resources spent on acquiring additional licenses would not make as much impact on a center's performance score as encouraging more students towards theses or dissertation topics related to IUCRC research projects.

Experts were not surprised that student topics contributed a high degree towards student involvement. "*Students who are more involved typically have a personal motive and interest beyond the research project. It's the students who are willing to work at home, continuing to conduct research that are the most engaged.*" Some students working as research assistants participate in the center as more of a job.

Experts believe a significant role can be played by university graduates hired into the field and by new methods for accelerating and

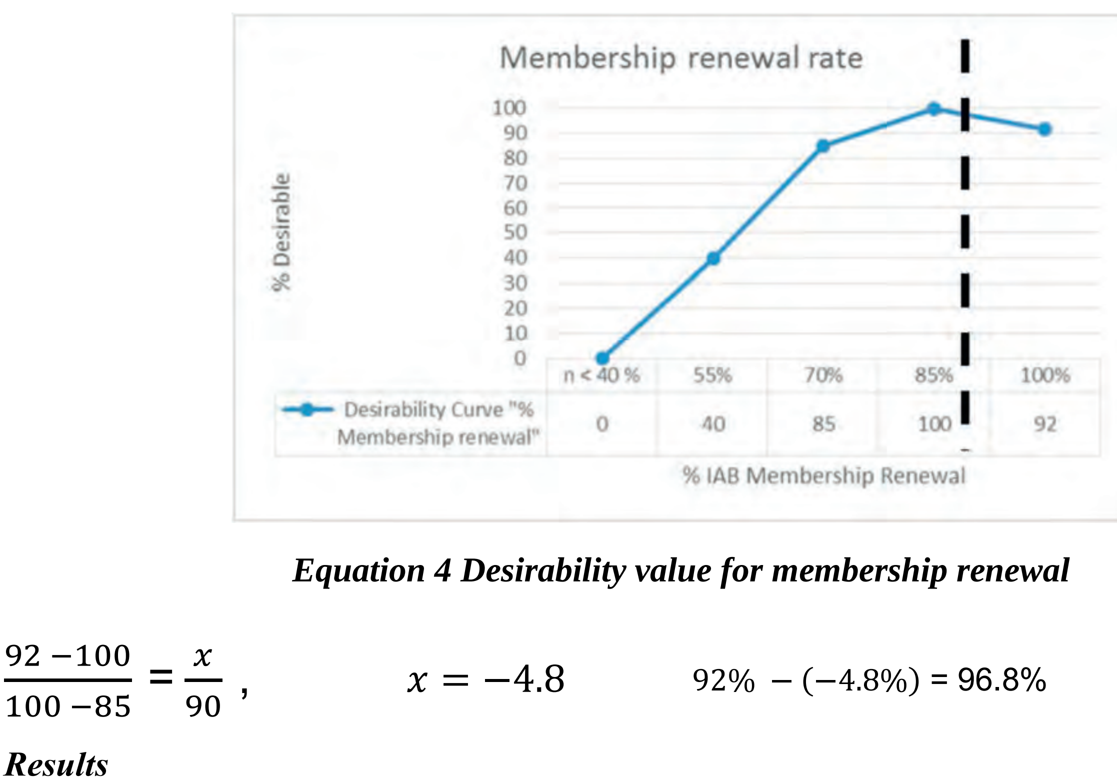


Fig. 10. WBC value for % membership renewal results.

Table 13
WBC metrics and desirability values.

Output decision element	Metric value (***m***,***jk***)	Desirability curve value d(m,jk)
Collaborative projects	0.33	0.28
Collaborative papers	0.00	0.00
IAB Member satisfaction	0.90	0.97
Leveraged funding	3.83	0.70
Researcher satisfaction	1.11	1.00
Student Mtg. attendance	0.52	0.73
Student research topic	0.00	0.03
Student research project	1.20	0.75
Student presentations	0.14	0.25
Student hires	2.00	1.00
Consulting	0.07	0.37
Papers published	0.63	0.80
Training and workshops	0.69	0.75
Shared resources	Both	1.00
New methods or processes	1.00	1.00
New products	0.00	0.50
Licenses	0.00	0.50

Table 14
Calculated performance evaluation score.

Output Contribution	Weights	*d(m,jk)*	Product
C. Research Projects	0.14	0.28	0.039
C. Research Papers	0.08	0.00	0.000
IAB Member Sat	0.06	0.97	0.058
Leveraged Funding	0.07	0.70	0.049
Researcher Sat.	0.04	1.00	0.040
Visiting Students	0.07	0.73	0.051
Student Topics	0.12	0.03	0.004
Student Projects	0.08	0.75	0.060
Student Presentations	0.05	0.25	0.012
Student Hires	0.06	1.00	0.060
Consulting	0.03	0.37	0.011
Papers	0.02	0.80	0.016
Training and Workshops	0.04	0.75	0.030
Shared Resources	0.03	1.00	0.030
New Methods/Proc.	0.07	1.00	0.070
New Products	0.02	0.50	0.010
Licenses	0.02	0.50	0.010
Sum of the Product			0.550

The shaded values represent the higher weighted decision elements. While this model has seventeen decision criteria notice how the top 2 account for 26% of the performance contribution. This means the decision criteria are not linearly related and that the method is able to separate more important elements from the ones that contribute towards the organizational performance to a lesser degree.

promoting knowledge and technology transfer. These results make sense because graduates have the opportunity to provide a long term impact to the field. This perspective was supported by the judgment provided by the expert panels reflected in student hires contributing approximately 6% towards the mission.

As shown, a strength of the model is that the more important decision criteria can be identified and their impact can be analyzed relatively quickly. This can be a powerful aid to managers and policy makers. However, what happens to the model when experts disagree about the decision criteria? This model and these values are subjective and not absolute. There are many reasons for differences. Some centers may have more difficulty with intellectual property issues because of their technology domain; they may instead focus on development of students. Efforts such as these could be diminished with this pure benchmarking approach.

The inconsistency and disagreement analysis provided new insights. For instance, one expert argued the fairness of one indicator: "*Inclusion of a metric for student hires may be problematic because there is a high percentage of International students.*" Therefore, some IUCRCs may have participating students who are legally not able to accept a position in a company if one was extended. They further qualified their argument expressing concern about possible screening practices that could be encouraged as a result of too much focus in this area. While the expert data uncovered some findings that may be of interest to policy makers and NSF IUCRC directors, a debate about the mission or objectives of the NSF program is beyond the scope of this research. Instead the goal here has been to measure the degree to which centers are meeting the mission as currently defined.

The metric measuring collaborative research also had a high amount of disagreement. Some experts advocated for only counting multi-site or multi-disciplined configured research teams, others stated that all were collaborative by definition because they had industry sponsorship. In general, all experts agree that "*collaborative projects are probably one area that has not been given enough focus*."

The Wood-Based Composites IUCRC was used to illustrate how a performance evaluation score is calculated using the model. One strength of the model is that decision criteria contributing to a higher degree towards the organizational performance can be readily identified. The case showed how improvement in outputs for the more heavily weighted decision elements could significantly improve performance.

The results and generalizability of the model was validate through consultation with experts. Experts expressed interest for a broadened study that examined how to make the model even more generalizable.

6. Conclusions

This research was able to successfully meet the original objectives set forth at the beginning of this paper. While this research was successful at taking a step towards closing the gaps identified in the literature, many still remain. Limitations included use of subjective data, development of proxy metrics and partial data sets. Future research opportunities are plentiful in this area including extensions to other

Table 15
Performance improvement recommendations.

Center	Pre-score	c_j	Suggested improvement	Contribution		New
				Current	Impact	Score
WBC	0.55	1	Encourage 1 student to select an IAB research project as their dissertation or thesis topic.	0	+0.05	0.65
		2	Projects 4/14 increase to 70%.	0.05	+0.05	

Table 16
Validation results.

Research validation	Test description for this research	Methods	Results
Content validity	The degree to which the content adequately describes the NSF IUCRC mission.	Delphi process during model development. Experts validated content and construct when 80% agreement was reached. Criteria and linked relationships were validated (Tran, 2013a).	Experts validated 17 of the decision criteria identified by literature.
Construct	Elements linked together creating the logic in a hierarchical construction.		Proxy metrics developed for several indicators for lack of data.
Criterion-related	Degree to which the criterion can capture the true value of the IUCRC's performance.	Expert review of case study analysis and results.	Experts were in general agreement with the results from the case study and determined the model is appropriate and generalizable.

NSF and NIH CRC and other types of CRC programs, methods for more robust sensitivity analysis, longitudinal studies to examine possible forecasting models for program sustainability and integration with proposal evaluation studies.

Increasingly important is the need for inter-disiplinary and inter-organizational collaborative research. Recognizing this need, the US National Science Foundation (NSF) has responded with funding and programmatic support for cooperative science and engineering research centers (CRCs). While evidence shows these centers are effective mechanisms for fundamental research, student development and knowledge and technology transfer; challenges remain to effectively measure and compare the performance of these organizations.

Organizational effectiveness is a difficult construct. Using the HDM, concepts were identified, validated by experts and linked together to construct a generalizable model. Transparency in how the decision variables impact the final performance scores was demonstrated by analyzing how a center could turn their performance upside-down by focusing on fewer than 20% of the outputs. Understanding where to shift resources can be a powerful decision aid to center directors. In one case example, it was demonstrated how the WBC center could obtain a significant performance increase by re-configuring project teams to include multi-disciplinary researchers and advising students towards completion of dissertation or theses using IAB projects as topics.

Centers were comparatively analyzed providing specific recommendations. The results were presented to an expert for criteria-related validity. The expert review validated the model and the results. The generalizability of the model was validated for the IUCRC program and interest was expressed for a broadened study to make the model even more generalizable.

6.1. Research contributions

This research begins to fill some of the gaps identified in literature. First, a system of outputs and metrics were presented from a balanced perspective. The hierarchical decision model (HDM) was introduced as a measurement system using both quantitative and qualitative metrics. The holistic study was validated using a 3-phased validation approach: 1) concept and content validation, 2) construct validation and 3) criterion-related validation. The criterion-related validity involved expert review of the results from a comparison of the performance of six case studies.

This research adds value to the field by offering a generalizable model and measurement system to compare performance of NSF science and engineering centers. It provides a new scoring method to compare and evaluate different IUCRCs. NSF center evaluators can then use these scores as a decision support tool for additional funding decisions and center managers can use these scores to analyze their portfolios in an objective, evidence-based manner increasing the achievement of their research objectives. The study effectively defined a set of output indicators painting a balanced-holistic picture of the NSF IUCRC program meeting the first objective of this research. While the generalizable model was only tested using the NSF IUCRC program, the model provides a new scoring method to compare and evaluate different IUCRCs in different programs.

A framework and metrics for evaluation was developed. Therefore, a new method for CRC performance comparison was introduced into the literature stream. This research begins to close the gap for cross CRC comparison by developing a generalizable model and a system for cross-center performance evaluation. The gap originally identified through literature was validated by experts. Gray agrees, "virtually all CRC outcome evaluation has been ad hoc, program-level evaluation studies" and that "these studies have tended to focus on technology transfer outcomes to industry"((Gray, 2008) p78).

6.2. Application contributions

The next contribution follows as a result of the first by disseminating the model and results of the study for improved assessment in the NSF IUCRC program. This study tested the model and the method by evaluating six (6) alternative IUCRCs. Many studies question if the traditional bibliometric indicators are the "right ones" and caution that they paint a "partial picture"(Wagner et al., 2011). The results of this research provide supporting evidence to this stream of literature by finding that new methods contribute significantly higher towards knowledge and technology transfer objectives than licenses. Federally funded CRCs are required to have transparency in their decision making processes. This research provides a new method that highlights disagreements helping to drive discussions and transparent decision making processes.

Representatives for the NSF SciPSI program remarked through an evaluation of this research agree that "*the need for understanding IUCRCs is important. They are a key policy lever used by the government to enhance translational research.*" *"Evaluating such centers remains difficult and often subjective, yet federal science agencies continue to invest considerable resources in them."* (NSF SciPHI program proposal evaluators).

This study benefits the research community by applying a flexible approach that combines qualitative and quantitative output indicators. Additional insight will be gained about the importance and use of output indicators. This holistic approach demonstrates a generalizable model that provides comparison among cooperative research centers.

Acknowledgements

This research is funded by Science of Science & Innovation Policy (SciSIP) program at the National Science Foundation (1632908: Doctoral Dissertation Research).

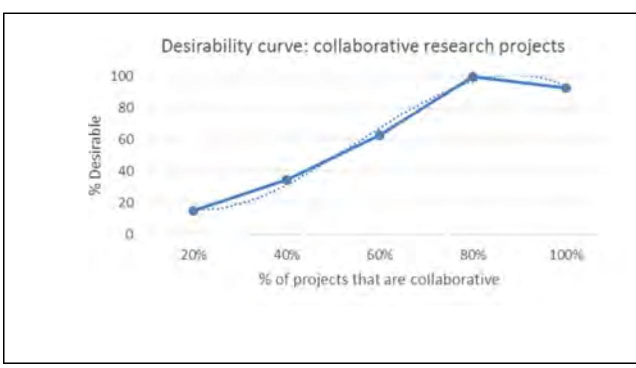
Appendix 1. Expert background

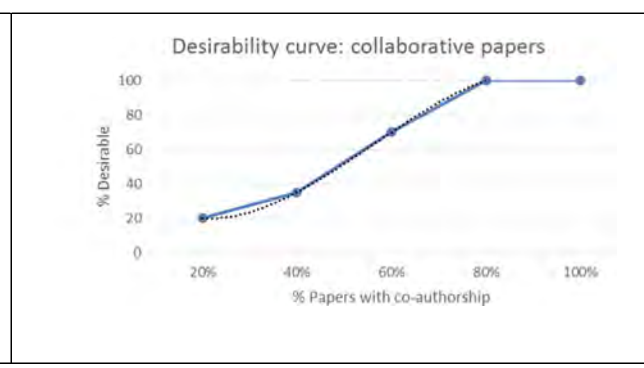
Expert #	Affiliation	Title	Background
1	NSF program	NSF program manager	NSF
2	North Carolina State U	Professor/director	C
3	Ohio State U	Professor	R
4	NSF IUCRC	NSF project manager	R
5	NSF consultant	Evaluator IUCRC	R
6	Iowa state	Assoc. prof/evaluator	NSF
7	University of Washington	Professor/evaluator	R
8	Palo Alto	Assoc. prof/evaluator	NSF
9	U of California, Berkeley	Professor/director	C
10	University of Florida	Evaluator IUCRC	NSF
11	Grand Valley State U	VP/evaluator	R
12	U of Colorado, Boulder	CoDirector IUCRC	C
13	North Carolina State U	Director IUCRC	C
14	SUNY Buffalo	Director IUCRC	C
15	Arizona State U	IUCRC staff	C
16	Univ of Buffalo	Director IUCRC	C
17	U of Arizona	Director IUCRC	C
18	Virginia Tech	Director IUCRC	C
19	Oregon State Univ	Director IUCRC	C
20	Virginia Tech	Director IUCRC	C
21	UC Davis	Director IUCRC	C
22	Brigham Young Univ	Director IUCRC	C
23	U of Arkansas	Director IUCRC	C
24	U of Tennessee, Knoxville	CoDirector IUCRC	C
25	U of California, Santa-Cruz	Center executive	C
26	Boise State	Evaluator IUCRC	NSF
27	North Texas	Professor/Assoc Director	C
28	Georgetown U	Director IUCRC	C
29	U of Washington	Director IUCRC	C
30	University of Tennessee	Evaluator IUCRC	NSF
31	Purdue U	Professor/evaluator	NSF
32	George Washington U	Professor	R
33	Arizona State U	Professor	R
34	Univ of Georgia	Evaluator IUCRC	NSF
35	Clarkson U	Director IUCRC	C
36	Purdue U	Ass Prof/Assist. Dir	C

Appendix 2. Desirability curves

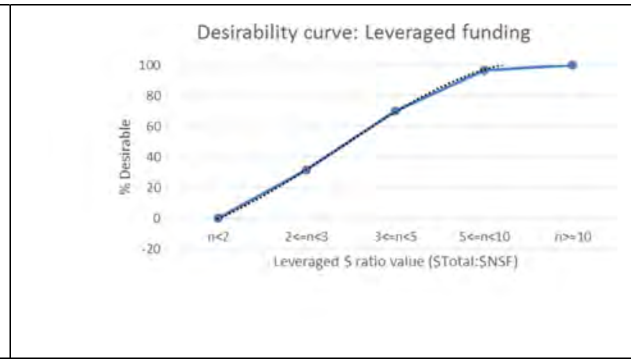
Metrics and desirability curves are presented relative to each of the six goals. Figures below show the respective desirability curves.

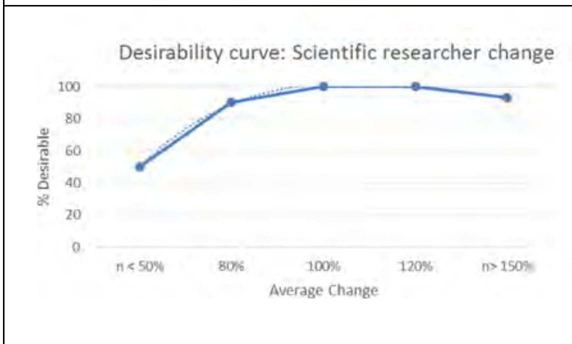
Desirability curves for new knowledge outputs





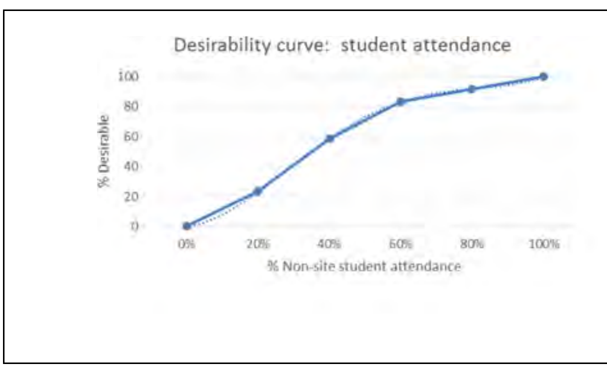
Desirability curves for stakeholder satisfaction outputs

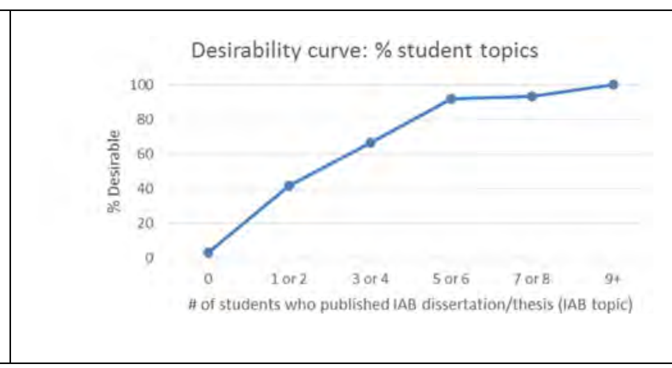




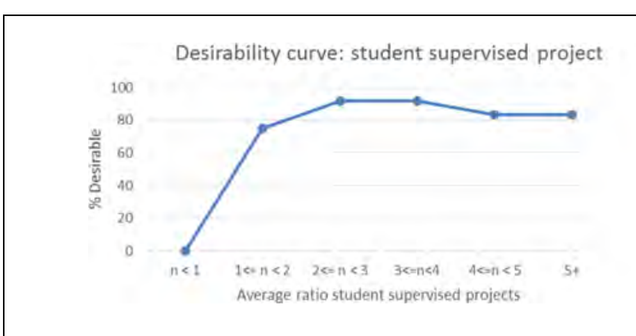
Some amount of stability for researchers is desired. Significant increases or decreases can be disruptive.

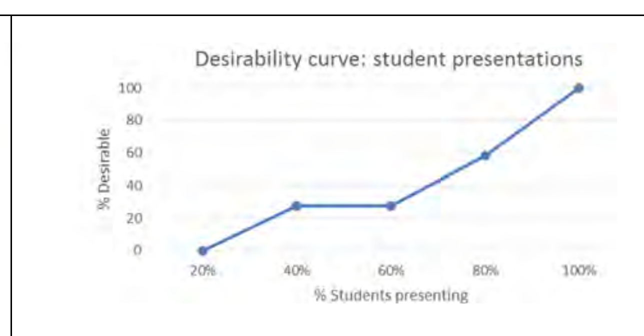
Desirability curves for student involvement outputs



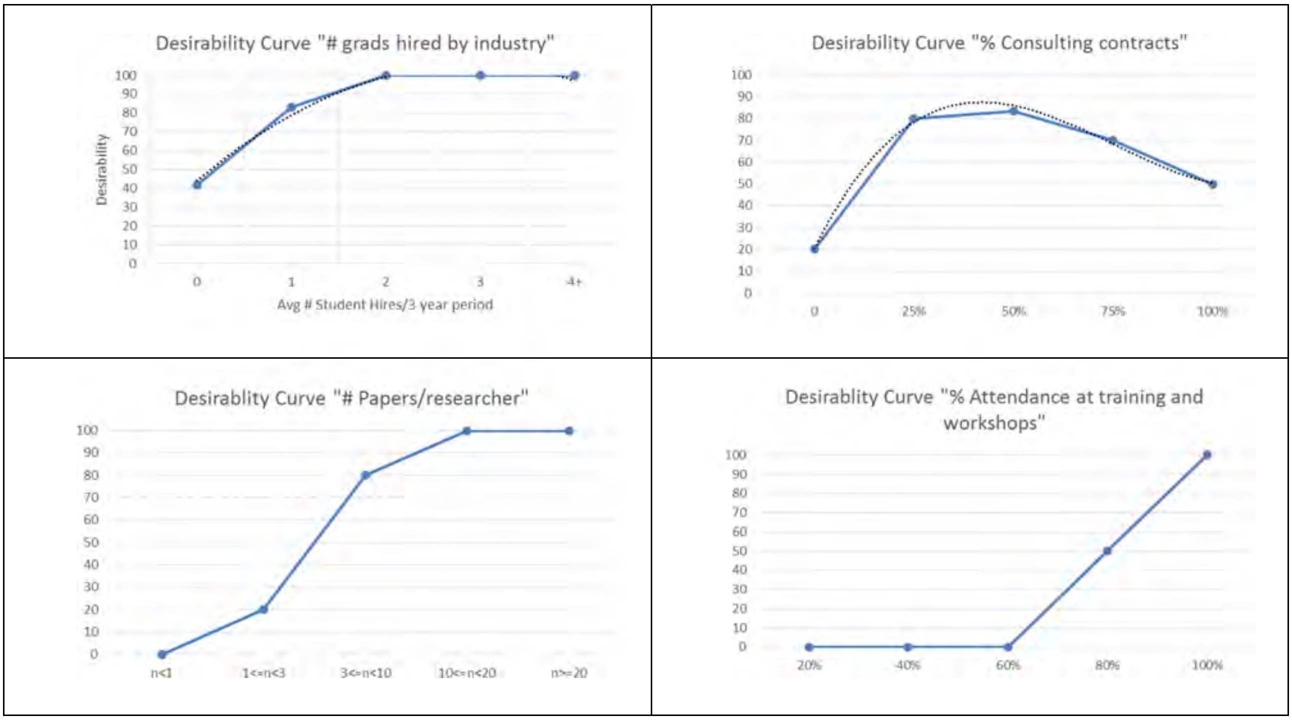


Desirability curves for student development outputs

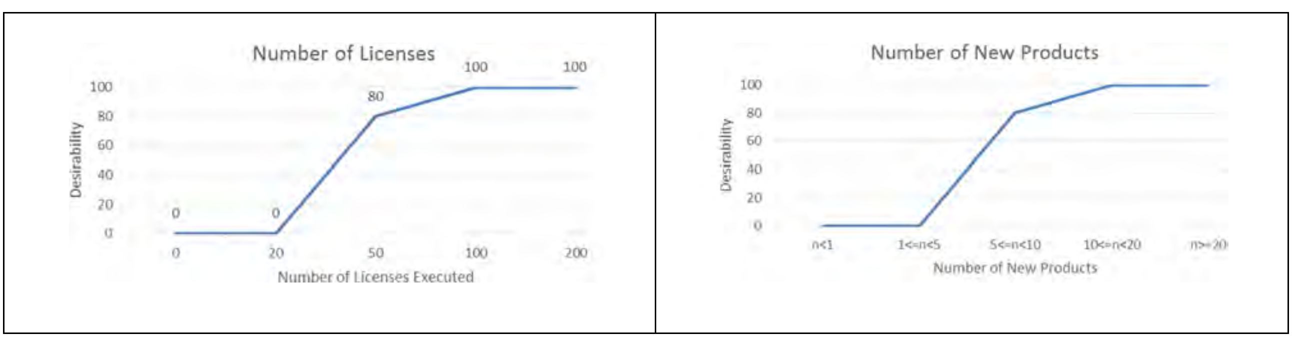




Desirability curves for KTT media



Desirability curves for KTT object outputs



Appendix 3. Additional center analyses

Center	Pre-score	c_j	Suggested improvement	Contribution		New
				Current	Impact	Score
Ma^2JIC	0.68	1	Increase multi-site/multi-discipline research project configurations from 7 to 14 of 25.	0.05	+0.04	0.76
		2	Increase co-publications from 5 to 9 of 15.	0.02	+0.04	
CPD	0.56	1	Increase multi-site/multi discipline research projects from 0 to 5 of 12.	0.02	+0.06	0.64
		2	Support student interest in selecting IUCRC topics for dissertation or thesis by 2 students	0.06	+0.02	
S^2ERC	0.57	1	Increase multi-site/multi-discipline research project configurations. Currently with 0 of 22 they should increase to 50% multi-site or multi-disciplined research project teams.	0.02	+0.06	0.63
CSR	0.46	1	Encourage 1 student to select an IAB research project as their dissertation or thesis topic.	0	+0.05	0.58
		2	Increase collaborative configuration from 0 to 6 of nine projects. Increase to 60%.	0.02	+0.06	
WEP	0.46	1	Encourage 1 student to select an IAB research project as their dissertation or thesis topic.	0	+0.05	0.57
		2	Increase collaborative configuration from 0 to 0 of 11 projects. Increase to 60%.	0.02	+0.06	
WBC	0.55	1	Encourage 1 student to select an IAB research project as their dissertation or thesis topic.	0	+0.05	0.65
		2	Projects 4/14 increase to 70%.	0.05	+0.05	

References

Chavez, V. (Vic), Feb. 2013. Creative enterprise as a solution vector for twenty first century problems. Technol. Forecast. Soc. Change 80 (2), 191–193.

NSF, "Industry & University Cooperative Research Program (I/UCRC)," *NSF 13-594*, 2013. [Online]. Available: http://www.nsf.gov/pubs/2013/nsf13594/nsf13594.htm. [Accessed: 16-Feb-2015].

White House Office of Management and Budget, "Science and Technology Priorities for the FY 2011 Budget," *Memoranda 2009: M-09-27*, 2011. [Online]. Available: www.whitehouse.gov/omb/memoranda_2009. [Accessed: 16-Feb-2015].

Roessner, "Outcomes and Impacts of the State/Industry- University Cooperative Research Centers (S/Iucrc) Program," 2000.

A. Abbasi and J. Altmann, "On the Correlation between Research Performance and Social Network Analysis Measures Applied to Research Collaboration Networks," 2011 *44th Hawaii Int. Conf. Syst. Sci.*, pp. 1–10, 2011.

Abbasi, A., Chung, K.S.K., Hossain, L., 2012. Egocentric analysis of co-authorship network structure, position and performance. Inf. Process. Manag. 48 (4), 671–679.

Abbasi, A., Wigand, R.T., Hossain, L., 2014. Measuring social capital through network analysis and its influence on individual performance. Libr. Inf. Sci. Res. 36 (1), 66–73.

Adams, J.D., Chiang, E.P., Starkey, K., 2001. Industry-university cooperative research centers. J. Technol. Transf. 26, 73–86.

Adner, R., Kapoor, R., 2010. Value creation in innovation ecosystems: how the structure of technological interdependence affects firm performance in new technology generations. Strateg. Manag. J. 31, 306–333.

Alkemade, F., Suurs, R.A.A., 2012. Patterns of expectations for emerging sustainable technologies. Technol. Forecast. Soc. Change 79 (3), 448–456 Mar.

Anderson, T.R., Daim, T.U., Lavoie, F.F., 2007a. Measuring the efficiency of university technology transfer. Technovation 27 (5), 306–318 May.

Anderson, T.R., Daim, T.U., Lavoie, F.F., 2007b. Measuring the efficiency of university technology transfer. Technovation 27 (5), 306–318 May.

Balconi, M., Laboranti, A., 2006. University-industry interactions in applied research: the case of microelectronics. Res. Policy 35, 1616–1630.

Behrens, T.R., Gray, D.O., 2001. Unintended consequences of cooperative research: impact of industry sponsorship on climate for academic freedom and other graduate student outcome. Res. Policy 30, 179–199.

C. Boardman and B. Bozeman, "Academic faculty as intellectual property in university-industry research alliances," *Econ. Innov. New Technol.*, no. 2, pp. 1–18, 2015.

Boardman, P.C., Corley, E.A., 2008. University research centers and the composition of research collaborations. Res. Policy 37 (2), 900–913.

Boardman, C., Gray, D., 2010. The new science and engineering management: cooperative research centers as government policies, industry strategies, and organizations. J. Technol. Transf. 35, 445–459.

Boardman, C., Ponomariov, B., 2011. A preliminary assessment of the potential for 'team science' in DOE Energy Innovation Hubs and Energy Frontier Research Centers. Energy Policy 39 (6), 3033–3035.

Bozeman, B., 2000. Technology transfer and public policy: a review of research and theory. Res. Policy 29, 627–655.

Bozeman, B., Boardman, C., 2013. Academic faculty in university research centers: neither capitalism's slaves nor teaching fugitives. J. High. Educ. 84 (1), 88–120.

Bozeman, B., Fay, D., Slade, C.P., 2013a. Research collaboration in universities and academic entrepreneurship: the-state-of-the-art. J. Technol. Transf. 38, 1–67.

Bozeman, B., Fay, D., Slade, C.P., 2013b. Research collaboration in universities and academic entrepreneurship: the-state-of-the-art. J. Technol. Transf. 38 (1), 1–67 Feb.

Bruneel, J., D'Este, P., Salter, A., 2010. Investigating the factors that diminish the barriers to university-industry collaboration. Res. Policy 39 (7), 858–868.

K. Cameron, "Measuring Organizational Effectiveness in Institutions of Higher Education," *Adm. Sci. Q.*, vol. 23, no. December, pp. 604–632, 1978.

Carayannis, E.G., von Zedtwitz, M., 2005. Architecting gloCal (global–local), real-virtual incubator networks (G-RVINs) as catalysts and accelerators of entrepreneurship in transitioning and developing economies: lessons learned and best practices from current development and business incubation. Technovation 25 (2), 95–110 Feb.

Carayannis, E., Del, M., Maria, G., Della, R., Carayannis, E., Del, M., Maria, G., Della, R., Villasalero, M., Campanella, F., Rosaria, M., Peruta, D., Del Giudice, M., Villasalero, M., 2014a. Managing the intellectual capital within government-university-industry R&D partnerships. J. Intellect. Cap. 15 (4), 611–630.

Carayannis, E., Del, M., Maria, G., Della, R., Carayannis, E., Del, M., Maria, G., Della, R., Villasalero, M., Campanella, F., Rosaria, M., Peruta, D., Del Giudice, M., Villasalero, M., 2014b. Managing the intellectual capital within government-university-industry R&D partnerships. J. Intellect. Cap. 15 (4), 611–630.

Chai, S., Shih, W., 2016. Bridging science and technology through academic–industry partnerships. Res. Policy 45 (1), 148–158.

Chen, H., Kocaoglu, D.F., 2008a. A sensitivity analysis algorithm for hierarchical decision models. Eur. J. Oper. Res. 185, 266–288.

Chen, H., Kocaoglu, D.F., 2008b. A sensitivity analysis algorithm for hierarchical decision models. Eur. J. Oper. Res. 185 (1), 266–288 Feb.

Chukumba, C., Jensen, R., 2005. University Invention, Entrepreneurship, and Start-Ups. Natl. Bur. Econ. Res.

Cleland, D.I., Kocaoglu, D.F., 1981. Engineering Management. McGraw-Hill, New York.

W. Cohen, R. Florida, and W. Goe, "University-Industry Research Centers in the United States." Center for Economic Development, H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, 1994.

W. M. Cohen, R. R. Nelson, and J. P. Walsh, "Links and Impacts: The Influence of Public Research on Industrial R&D," *Manag. Sci.*, vol. 48, no. June 2015, pp. 1–23, 2002.

Comacchio, A., Bonesso, S., Pizzi, C., 2011. Boundary spanning between industry and university: the role of Technology Transfer Centres. J. Technol. Transf. 37 (6), 943–966 Aug.

Corley, E.A., Boardman, P.C., Bozeman, B., 2006. Design and the management of multi-institutional research collaborations: theoretical implications from two case studies. Res. Policy 35, 975–993.

M. H. Cragin, L. Nichols, M. Simon, and S. M. Watts, "Measuring science: Emerging tools for analysis of federal R&D investments," *Proc. ASIST Annu. Meet.*, vol. 49, no. 1, 2012.

Cunningham, J., O'Reilly, P., O'Kane, C., Mangematin, V., 2014. The inhibiting factors that principal investigators experience in leading publicly funded research. J. Technol. Transf. 39, 93–110.

T. U. Daim, G. Rueda, H. Martin, and P. Gerdsri, "Forecasting emerging technologies: use of bibliometrics and patent analysis," Technol. Forecast. Soc. Change, vol. 73, no. 8, pp. 981–1012, Oct. 2006.

Devine, M., James, T., Adams, T., 1987. Government supported industry-university research centers: issues for successful technology transfer. J. Technol. Transf. 12 (1), 27–37.

Ecken, P., Gnatzy, T., von der Gracht, H., 2011. Desirability bias in foresight: consequences for decision quality based on Delphi results. Technol. Forecast. Soc. Change 78 (9), 1654–1670 Nov.

Elkarmi, F., Mustafa, I., 1993. Increasing the utilization of solar energy technologies (SET) in Jordan. Energy Policy 21 (9), 978–984 Sep.

Etzkowitz, H., 1998. The norms of entrepreneurial science: cognitive effects of the new university–industry linkages. Res. Policy 27, 823–833.

Etzkowitz, H., Leydesdorff, L., 2000a. The dynamics of innovation: from National Systems and 'Mode 2' to a Triple Helix of university–industry–government relations. Res. Policy 29 (2), 109–123 Feb.

Etzkowitz, H., Leydesdorff, L., 2000b. The dynamics of innovation: from National Systems and 'Mode 2' to a Triple Helix of university–industry–government relations. Res. Policy 29 (2), 109–123 Feb.

Feldman, M.P., Kelley, M.R., 2006a. The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. Res. Policy 35, 1509–1521.

Feldman, M.P., Kelley, M.R., 2006b. The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. Res. Policy 35, 1509–1521.

Freeman, C., Soete, L., 2009. Developing science, technology and innovation indicators: what we can learn from the past. Res. Policy 38 (4), 583–589 May.

Garner, J.G., Porter, A.L., Newman, N.C., Crowl, T.A., 2012. Assessing research network and disciplinary engagement changes induced by an NSF program. Res. Eval. 21 (2), 89–104.

Gaughan, M., Corley, E., 2010. Science faculty at US research universities: the impacts of university research center-affiliation and gender on industrial activities. Technovation 30 (3), 215–222.

F. W. Geels, "Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective," Res. Policy, vol. 39, no. 4, pp. 495–510, May 2010.

Geisler, E., 2003. Benchmarking inter-organisational technology cooperation: the link between infrastructure and sustained performance. Int. J. Technol. Manag. 25 (8), 675–702.

Geisler, E., 2010. Measuring the contributions of knowledge management systems to the strategic competitiveness of organizations: a review and a model. Int. J. Innov. Technol. Manag. 07 (02), 89–107.

Georghiou, L., Cassingena Harper, J., 2013. Rising to the challenges—reflections on future-oriented technology analysis. Technol. Forecast. Soc. Change 80 (3), 467–470 Mar.

P. Gerdsri and D. F. Kocaoglu, "HDM for developing national emerging technology strategy and policy supporting sustainable economy: A case study of nanotechnology for Thailand’s agriculture," in *PICMET '08* - 2008 *Portland International Conference on Management of Engineering & Technology*, Cape Town, South Africa, 2008, pp. 344–350.

Geuna, A., Muscio, A., 2009. The governance of university knowledge transfer: a critical review of the literature. Minerva 47, 93–114.

E. Gibson, "Supporting Sustainable technology cluster development: a performance measurement problem," in *PICMET* 2015, 2015, pp. 1–16.

Graham, R., 2013. Technology Innovation Ecosystem Benchmarking Study: Key Findings from Phase 1.

D. O. Gray, "Making team science better: Applying improvement-oriented evaluation principles to evaluation of coorperative research centers," *New Dir. Eval.*, no. 118, pp. 107–114, 2008.

Gray, D.O., Steenhuis, H.J., 2003. Quantifying the benefits of participating in an industry university research center: an examination of research cost avoidance. Scientometrics 58 (2), 281–300.

Gray, D.O., Walters, S.G., 1998. Managing the Industry/University Cooperative Research Center: A Guide for Directors and Other Stakeholders. Battelle Press, Columbus.

Gray, D., Lindblad, M., Rudolph, J., 2001. Industry–university research centers: a multivariate analysis of member retention. J. Technol. Transf. 26, 247–254.

Gray, D., Sundstrom, E., Tornatzky, L.G., McGowen, L., 2011. When Triple Helix unravels: a multi-case analysis of failures in industry–university cooperative research centres. Ind. High. Educ. 25 (5), 333–345.

Gray, D.O., Rivers, D., Vermont, G., 2012a. Measuring the Economic Impacts of the NSF Industry/University Cooperative Research Centers Program: A Feasibility Study. Arlington, Virginia.

D. O. Gray, L. Tornatzky, L. McGowen, and E. Sundrome, "Research Center Sustainability and Survival: Case Studies of Fidelity, Reinvention and Leadership of Industry/University Cooperative Research Centers," vol. 0631414, p. 93, 2012b.

Gray, D.O., Boardman, C., Rivers, D., 2013. The new science and engineering management: cooperative research centers as intermediary organizations for government

policies and industry strategies. In: Cooperative research centers and technical innovation. Springer, New York, pp. 3–33.
Gray, D.O., Leonchuk, O., McGowen, L.C., Michaelis, T., 2014a. 2012–2013 Process Outcome Survey Results Descriptive Statistics Compiled from Industry and Faculty Surveys. North Carolina State University, Raleigh.
Gray, D.O., Leonchuk, O., McGowen, L.C., Michaelis, T., 2014b. 2012–2013 Process Outcome Survey Results Descriptive Statistics Compiled from Industry and Faculty Surveys. North Carolina State University, Raleigh.
Gray, D., Rivers, D., McGowen, L., Leonchuck, L., Michaelis, T., 2014c. SBIR/STTR Supplemental Opportunity in I/UCRCs. Arlington, Virginia.
Grupp, H., Linstone, H.A., 1999. Around the globe resurrection and new paradigms. Technol. Forecast. Soc. Change 94, 85–94.
Hall, B., Link, A., Scott, J., 2003. The review of economics and statistics. Rev. Econ. Stat. 85 (2), 485–491.
Hämäläinen, R.P., Karjalainen, R., 1992. Decision support for risk analysis in energy policy. Eur. J. Oper. Res. 56 (2), 172–183 Jan.
Hayton, J.C., Sehili, S., Scarpello, V., 2010. Why do firms join consortial research centers? An empirical examination of firm, industry and environmental antecedents. J. Technol. Transf. 35, 494–510.
Hicks, D., 2012. Performance-based university research funding systems. Res. Policy 41 (2), 251–261.
Iskin, I., 2014. An Assessment Model for Energy Efficiency Program Planning in Electric Utilities: Case of Northwest U.S. Portland State University.
Katz, J.S., Martin, B.R., 1997. What is research collaboration? Res. Policy 26 (1), 1–18.
Kostoff, R.N., 2005. Science and Technology Metrics. Arlington, Virginia.
Lee, Y.S., 2000. The sustainability of university-industry research collaboration. J. Technol. Transf. 25, 111–133.
Lee, S.K., Yoon, Y.J., Kim, J.W., 2007. A study on making a long-term improvement in the national energy efficiency and GHG control plans by the AHP approach. Energy Policy 35 (5), 2862–2868 May.
Lee, S.K., Mogi, G., Kim, J.W., 2008. The competitiveness of Korea as a developer of hydrogen energy technology: the AHP approach. Energy Policy 36, 1284–1291.
Leydesdorff, L., Carley, S., Rafols, I., 2013. Global maps of science based on the new Web-of-Science categories. Scientometrics 94 (2), 589–593.
Li, E.Y., Liao, C.H., Yen, H.R., 2013. Co-authorship networks and research impact: a social capital perspective. Res. Policy 42 (9), 1515–1530.
Link, A.N., Siegel, D.S., 2005a. University-based technology initiatives: quantitative and qualitative evidence. Res. Policy 34 (3), 253–257 Apr.
Link, A.N., Siegel, D.S., 2005b. University-based technology initiatives: quantitative and qualitative evidence. Res. Policy 34 (3), 253–257.
B. Martin and R. Johnston, "Technology Foresight for Wiring Up the National Innovation System Experiences in Britain, Australia, and New Zealand," Technol. Forecast. Soc. Change, vol. 60, no. 1999, pp. 37–54, 1999.
Masum, H., Ranck, J., Singer, P.A., 2010. Five promising methods for health foresight. Foresight 12 (1), 54–66 Feb.
McGowan, L., 2012. Program Sustainability for Cooperative Research Centers: A Longitudinal Analysis. North Carolina State University.
Y. Motoyama, J. Konczal, J. Bell-masterson, and A. Morelix, "Think Locally, Act Locally: Building a Robust Entrepreneurial Ecosystem," 2014a.
Y. Motoyama, J. Konczal, J. Bell-masterson, and A. Morelix, "Think Locally, Act Locally: Building a Robust Entrepreneurial Ecosystem," 2014b.
J. L. Mumpower and T. R. Stewart, "Expert Judgement and Expert Disagreement," *Think. Reason.*, vol. 2, no. February 2015, pp. 191–212, 1996.
Palomares-Montero, D., Garcia-Aracil, A., 2011. What are the key indicators for evaluating the activities of universities? Res. Eval. 20 (5), 353–363.
Penfield, T., Baker, M.J., Scoble, R., Wykes, M.C., 2014. Assessment, evaluations, and definitions of research impact: a review. Res. Eval. 23 (1), 21–32.
Perkmann, M., Walsh, K., 2007. University-industry relationships and open innovation: towards a research agenda. Int. J. Manag. Rev. 9 (4), 259–280.
Perkmann, M., Neely, A., Walsh, K., 2011a. How should firms evaluate success in university-industry alliances? A performance measurement system. R Manag. 41 (2), 202–216.
Perkmann, M., Neely, A., Walsh, K., 2011b. How should firms evaluate success in university-industry alliances? A performance measurement system: how should firms evaluate success in university-industry alliances? RD Manag. 41 (2), 202–216 Mar.
Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Krabel, S., Kitson, M., Llerena, P., Lissoni, F., Salter, A., Sobrero, M., 2013. Academic engagement and commercialisation: a review of the literature on university-industry relations. Res. Policy 42 (2), 423–442.
Phan, K., 2013. Innovation Measurement: A Decision Framework to Determine Innovativeness of a Company. Portland State University.
Phan, P.H., Siegel, D.S., 2006. The effectiveness of university technology transfer. Found. Trends® Entrep. 2 (2), 77–144.
E. Piva and C. Rossi-Lamastra, "Systems of indicators to evaluate the performance of university-industry alliances: a review of the literature and directions for future research," *Meas. Bus. Excell.*, vol. 17, no. June, pp. 40–54, 2013.
Ponomariov, B.L., Boardman, P.C., 2010. Influencing scientists' collaboration and productivity patterns through new institutions: university research centers and scientific and technical human capital. Res. Policy 39 (5), 613–624.
Porter, A.L., Rafols, I., 2009. Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics 81 (3), 719–745.
Quinn, R.E., Rohrbaugh, J., 2014. A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis. Manag. Sci. 29 (3), 363–377.
Quinn, R.E., Rohrbaugh, J., Rohrbaugh, J., 1981. A competing values approach to organizational effectiveness. Public Product. Rev. 5 (2), 122–140.
Rafols, I., Porter, A.L., Leydesdorff, L., 2009. Science overlay maps: a new tool for research policy and library management. J. Am. Soc. Inf. Sci. Technol. 61 (9), 1871–1887.
Ramanathan, J., Ramnath, R., Herold, M.J., Wierwille, B.J.R., 2013. An agile translation process for complex innovations: an industry/university cooperative research center case study. Proc. - Front. Educ. Conf. FIE 1532–1538.
D. Rivers, "To Join or Not to Join: Individual and Sub-organizational Factors Affecting Industry Membership in University-based Cooperative Research Centers," *8th Int.* Triple Helix Conf., 2010.
Robinson, D.K.R., Huang, L., Guo, Y., Porter, A.L., 2013. Forecasting Innovation Pathways (FIP) for new and emerging science and technologies. Technol. Forecast. Soc. Change 80 (2), 267–285 Feb.
Roessner, J.D., 1998. Review of Progress in quantitative nondestructive evaluation. Rev. Prog. Quant. Nondestruct. Eval. 12, 31–39.
Roessner, D., Manrique, L., Park, J., 2010. The economic impact of engineering research centers: preliminary results of a pilot study. J. Technol. Transf. 35, 475–493.
Rogers, J.D., Bozeman, B., 1997. Basic research and the success of federal lab-industry partnerships. J. Technol. Transf. 22 (3), 37–47.
Rohrbeck, R., 2010. Harnessing a network of experts for competitive advantage: technology scouting in the ICT industry. RD Manag. 40 (2), 169–180 Mar.
Rohrbeck, R., Kaab, S., 2013. Collaborative business modelling for systemic and sustainability innovations. Int. J. Technol. Manag. 63 (1/2), 4–23.
R. Rohrbeck, L. Konnertz, and S. Knab, "Collaborative business modelling for systemic and sustainability innovations," *Int. J. Technol. Manag.*, vol. 63, no. 1/2, p. 4, 2013.
Ruegg, R., 2006. Bridging from Project Case Study to Portfolio Analysis in a Public R & D Program, NIST GCR 06-891. Gaithersburg, MD.
Ruegg, R., Feller, I., 2003. A toolkit for evaluating public R&D investment. Gaithersburd, MD: National Institute of Standards and Technology, NIST GCR 03-857.
T. L. Saaty, "Decision making with the analytic hierarchy process," *Int. J. Serv. Sci.*, vol. 1, no. 1, p. 83, 2008.
Schmoch, U., Schubert, T., Jansen, D., Heidler, R., von Görtz, R., 2010. How to use indicators to measure scientific performance: a balanced approach. Res. Eval. 19 (1), 2–18.
L. I. Schultz, "Measuring the Activities of Collaborative Research Centers," *Proc. 17th Int.* Conf. Sci. Technol. Indic., vol. 2, pp. 743–755, 2012.
Scott, C.S., 2014. Compendium 2014. NSF, Raleigh, N.C.
Shane, S., Somaya, D., 2007a. The effects of patent litigation on university licensing efforts. J. Econ. Behav. Organ. 63 (4), 739–755 Aug.
Shane, S., Somaya, D., 2007b. The effects of patent litigation on university licensing efforts. J. Econ. Behav. Organ. 63, 739–755.
Siegel, D.S., Waldman, D.A., Atwater, L.E., Link, A.N., 2004. Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: qualitative evidence from the commercialization of university technologies. J. Eng. Technol. Manag. 21 (1–2), 115–142.
R. W. Smilor, D. V Gibson, and G. Kozmetsky, "Creating the Technopolis: High Technology Development in Austin, Texas," *Bus. Ventur.*, pp. 49–67, 1989.
Steers, R.M., 1975. Problems in the measurement of organizational effectiveness. Adm. Sci. Sci. 10 (4), 546–558.
E. Sundstrom and D. O. Gray, "Fostering Team Science: Innovative leadership practices in NSF Industry/University Research Cooperative Centers," in *Triple Helix VIII: International Conference on Government, University & Industry Linkages*, 2010, no. October, pp. 20–22.
Tran, T.A., 2013a. Strategic Evaluation of University Knowledge and Technology Transfer Effectiveness. Portland State University.
Tran, T.A., 2013b. Strategic Evaluation of University Knowledge and Technology Transfer Effectiveness. Portland State University.
Tsai, W., 2015. Knowledge transfer in intraorganizational networks: effects of network position and absorptive capacity on business unit innovation and performance. Acad. Manag. J. 44 (5), 996–1004.
Tyler, J.E., 2013. Redeploying Bayh-Dole: beyond merely doing good to optimizing the potential in results of taxpayer-funded research. J. Technol. Transf. 38, 911–929.
Wagner, C.S., Roessner, J.D., Bobb, K., Klein, J.T., Boyack, K.W., Keyton, J., Rafols, I., Börner, K., 2011. Approaches to understanding and measuring interdisciplinary scientific research (IDR): a review of the literature. J. Inf. Secur. 5, 14–26.
J. Wallenius, J. S. Dyer, P. C. Fishburn, R. E. Steuer, S. Zionts, K. Deb, J. S. Dyer, P. C. Fishburn, and R. E. Steuer, "Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead," *Manage. Sci.*, no. June 2016, 2008.
Wallner, T., Menrad, M., 2004. Extending the innovation ecosystem. Proc. IXXII ISPIM Conf. Hamburd Ger. 1–9.
Yokoo, Y., Okuwada, K., 2013. Identifying expected areas of future innovation by combining foresight outputs. Foresight 15 (1), 6–18.