Extending the Theory of Goal Ambiguity to Programs: Examining the Relationship between Goal Ambiguity and Performance

One of the main assumptions of empirical studies conducted on the influence of goal ambiguity in public management is that goal ambiguity relates negatively to performance. However, this relationship has rarely been tested at the program level because common goal ambiguity and performance measures for disparate government programs have been scant. The availability of Program Assessment Rating Tool (PART) results for a number of federal programs provides the opportunity for an analysis testing the foregoing assumption. Measures of program goal ambiguity—target, timeline, and program evaluation—are shown to have negative relationships with different program performance scores, taking into account alternative influences or biases on performance. This analysis extends the theory of goal ambiguity by providing the first analysis of large-sample federal programs. The theoretical and practical implications are presented in the discussion and conclusion.

On the other hand, evaluating the performance of governmental activities, a major topic in public administration, has received increased attention in recent years (Heinrich 2007; Lewis 2008; Meier and O’Toole 2007; Moynihan 2008; Radin 2008; Walker and Boyne 2006). As part of this trend, the U.S. Office of Management and Budget (OMB) created the Performance Assessment Rating Tool (PART) and administered it from 2004 to 2010 to assess the performance of federal programs. PART assessed four dimensions of administrative performance—design, planning, management, and results—and provided an overall performance rating that combined these four categories. This article considers these PART performance scores for the goal ambiguity–performance model, as discussed later. Common performance measures for disparate government programs are rare, especially in the United States at the federal level, and the PART scores provide performance indicators comparable across programs and agencies.

The availability of the PART assessment results for a number of federal programs (188 programs for which the PART performance score reassessment records could be tracked between 2004 and 2008, for the present analysis) provided an opportunity to analyze the relationship between goal ambiguity and performance at the program level. This goal ambiguity–performance model at the program level differs from previous organization-level models in terms of goal ambiguity concepts, ways of measuring the concepts, and control variables that can influence performance. At the same time, this article replicates evaluative goal ambiguity from Chun and Rainey’s (2005b) agency-level goal ambiguity research at the program level. In these ways, this analysis contributes to enhancing the generalizability of goal ambiguity theory or extending the theory to the program level in public management. The measures of program goal ambiguity—target, timeline, and program evaluation—for the current analysis show strong and consistent negative relationships with different PART scores. This article provides the first analysis of large-sample federal programs linking multiple goal ambiguity measures to multiple program performance dimensions and comparing different associations between them.

Goal-Setting Theory and Goal Ambiguity Research in Public Management

Goal-Setting Theory

Goal-setting theory premises that a goal can affect employees’ motivation and performance (Latham and Locke 1991; Lee, Locke, and Latham 1989). Regarding goal content, which indicates what needs
to be achieved, numerous empirical studies have demonstrated that when goals are specific (job-goal specificity) and reasonably difficult (goal difficulty), employee motivation and performance increase (e.g., Lee, Locke, and Latham 1989; Locke and Latham 1990, 2002). Most of these empirical studies used the individual as the unit of analysis. Smith, Locke, and Barry (1990) extended the unit of analysis from the individual to the organization, reporting that in an experiment with employees in private companies, setting specific organizational goals also contributed to an improvement in organizational performance. The reason is that employees in organizations that present specific organizational goals are more likely to take notice of goal discussions and to be committed to the goal (Smith, Locke, and Barry 1990).

More recently, another noticeable change in goal-setting literature has been the extension of the range of the organizational setting from schools or firms to government organizations. For example, particularly in relation to goal specificity, Wright (2007) reported that job-goal specificity had a direct effect as well as an indirect effect, through its influence on employee self-efficacy, on work motivation in a large New York State agency. These empirical studies testing goal-setting theory in public management have focused not on the agency level but only on the individual level.

**Naissance of a Theory of Goal Ambiguity in Public Management**

Consistent with goal-setting theory, especially as it relates to goal specificity, a long-lasting assertion of numerous scholars in public management and political science has been that organizational goals of government agencies are more ambiguous (goal ambiguity) than those of private businesses and that this goal ambiguity has important (mainly negative) consequences for government agencies and their employees (e.g., Allison 1983; Chun and Rainey 2005b; Dahl and Lindblom 1953; Downs 1967; Heinrich 1999; Loew 1979; Lynn 1981; Matland 1995; Rainey 1993; Ripley and Franklin 1982; Wildavsky 1979; Wilson 1989). The suggested main reasons for the goal ambiguity include the lack of clear profit indicators, political intervention from diverse stakeholders, and conflicting values (e.g., equity and efficiency) that government agencies should pursue (Rainey 2010). Further, some scholars noticed that ambiguous goals are needed to a certain degree to get leeway for political compromise with diverse interest groups and to sustain political support from them (Lindblom 1959; Stazyk and Goerdel 2011; Wildavsky 1979). Thus, ambiguous goals of government agencies have been recognized as an inevitable product of the political decision-making process.

This body of work has contributed to a theory of goal ambiguity in public management; in the article “Toward a Theory of Goal Ambiguity in Public Organizations,” Rainey (1993) suggested a comprehensive framework of goal ambiguity using the aforementioned main reasons as starting points. This framework contains numerous antecedents (e.g., public versus private distinction [lack of clear profit indicators], organizational autonomy, political power, influence of political authorities, organization size, manager’s goal clarification) and consequences (e.g., manager’s perceived goal ambiguity, employee work and public service motivation, employee commitment, organizational performance and accountability) of goal ambiguity in government agencies, with relevant propositions. Further, the possibility of goal ambiguity as a mediator or a moderator between the antecedents and consequences was also suggested. Rainey’s framework served as an important foundation for later empirical studies of goal ambiguity in public management; most of those studies, as discussed later, have provided evidence for the propositions suggested by Rainey.

The first stream of empirical studies (e.g., Lan and Rainey 1992; Rainey 1983; Rainey and Bozeman 2000) focused on the ubiquitous claim that organizational goals in government agencies are less clear or more ambiguous than those in private firms. They asked managers in both sectors to respond to questions about the clarity or ambiguity of their organization’s goals; however, little difference was found between the responses of the government and business managers. Such findings led Rainey and Bozeman (2000) to suggest methodological issues: first, the concept of “goal ambiguity” is ambiguous in itself; second, “Maybe the questionnaire items asking for simple, direct responses about the clarity and measurability of organizational goals are too simple to assess this dimension” (452). That is, these two scholars criticized the empirical studies depending on perceptual data and taking a one-dimensional approach to goal ambiguity (Jung 2011).

**Empirical Research on Antecedents of Organizational Goal Ambiguity in Public Management**

Starting with Chun and Rainey’s (2005b) empirical research, the last decade has witnessed a significant advance in empirical research on organizational goal ambiguity at the organizational and individual (manager) levels. Empirical studies on antecedents of organizational goal ambiguity that have extended the reasons for goal ambiguity can be divided into two lines. One line of research, the analysis unit of which is the government agency, follows from Chun and Rainey’s (2005b) new approach to conceptualize and measure organizational ambiguity. Unlike previous studies that used a survey and a single-dimension goal ambiguity, Chun and Rainey defined organizational goal ambiguity as “the extent to which an organizational goal or set of goals allows leeway for interpretation, when the organizational goal represents the desired future state of the organization” (2005b, 2) and developed objective measures of its four dimensions: mission comprehension ambiguity (Gunning fog index of an agency’s mission statement), directive goal ambiguity (ratio of the number of rules issued by an agency to the number of laws administered by the agency), evaluative goal ambiguity (percentage of output-oriented performance indicators [as opposed to outcome indicators] out of total performance indicators for an agency), and priority ambiguity (number of long-term goals and annual performance targets). Furthermore, they reported that various political environments or organizational characteristics, such as financial publicness, institutional location (independent versus within a cabinet department), competing demands, policy problem complexity, and regulatory responsibility, form the different levels of organizational goal ambiguity across government agencies. Later Lee, Rainey, and Chun (2009, 2010) reported that higher levels of political salience, including congressional salience, presidential salience, and media salience, and work complexity led to an increase in directive goal ambiguity as well as evaluative goal ambiguity.

The other line of research uses the survey questions from Phase II of the National Administrative Studies Project (NASP-II), which surveyed public managers, and measures a single-dimension
Empirical studies have focused broadly on two consequences of goal ambiguity in public management: organizational performance and employee perceptions of work attitudes and red tape. Regarding organizational consequences, the argument that "the clarity [or the specificity] with which [organizational] goals are expressed has implications for performance" has substantial overlap with the goal ambiguity literature (Stazyk and Goerdel 2011; see also Wright and Davis 2003). Empirical studies concerning the relationship between organizational goal ambiguity and performance can be divided into two categories: calculating objective measures of goal ambiguity and using survey items for organizational goal ambiguity. The former group includes Chun and Rainey (2005a) and Jung (2011). These studies, commonly based on the multidimensionality of goal ambiguity, empirically examined whether organizational performance can benefit from clear organizational goals (goal clarity). Chun and Rainey (2005a) linked the aforementioned objective dimensions of goal ambiguity to perceived organizational performance, including managerial effectiveness, customer service orientation, productivity, and work quality. They found that the directive and evaluative goal ambiguity measures related consistently and negatively to the perceived performance dimensions. More recently, Jung (2011) developed two new dimensions of goal ambiguity, target-specification ambiguity and time-specificity ambiguity (which were later renamed target ambiguity and timeline ambiguity), and replicated Chun and Rainey’s (2005b) priority ambiguity from the PART data. Then Jung (2011) showed that these three measures of goal ambiguity had a negative effect on actual goal attainment rates of U.S. federal agencies as a measure of organizational performance. Stazyk and Goerdel (2011) used the NASP-II survey but observed the same result, a negative relationship between one-dimensional goal ambiguity and perceived organizational effectiveness. Hence, regardless of the data source, the association between organizational goal ambiguity and performance has been consistently negative.

However, with anecdotal evidence, some scholars have suggested negative effects of clear goals and positive effects of ambiguous goals on performance. Contrary to the results of numerous goal-setting studies, a goal with a clear and narrowly focused target can degrade employee and organizational performance by leading employees to neglect important but unclear goals and by eroding the cooperative organizational culture (e.g., decreasing extrarole behavior) (Ordoñez et al. 2009; Wright et al. 1993). Moreover, as opposed to the long-standing assertion in public management regarding the adverse effects, ambiguous goals can help reduce interventions from external stakeholders and to increase managerial autonomy (Maynard-Moody and McClintock 1987); such goals can facilitate meaningful communication and discussion among stakeholders and managers and provide learning opportunities (Noordegraaf and Abma 2003).

On the other hand, some empirical studies have focused on work attitudes and red tape as consequences of goal ambiguity. These studies have consistently provided evidence of negative and positive relationships, respectively. Specifically, the objective or subjective dimensions of organizational goal ambiguity mentioned earlier have been found to be negatively associated with public service motivation (Jung and Rainey 2011), job satisfaction (Chun and Rainey 2006; Jung 2013b; Wright and Davis 2003), and organizational commitment (Hassan and Rohrbaugh 2011) but positively associated with turnover intention of public employees (Jung 2012b). Moreover, some authors (Chun and Rainey 2006; Pandey and Rainey 2006) reported that a higher level of goal ambiguity predicts a greater perception of red tape. As discussed in this section, most empirical studies concerning the consequences of organizational goal ambiguity have demonstrated its adverse effects, irrespective of the subjectivity and objectivity of goal ambiguity measures and the data sources.

In summary, empirical studies on goal ambiguity regarding both its antecedents and consequences have broadly diverged into two streams: those with objective measures and those with survey measures. These two streams have made important contributions to the refinement of the concept of goal ambiguity and demonstrates that the multiple dimensions have differing important antecedents and differing effects on consequences. The latter stream, with a one-dimensional goal ambiguity measure, adds different antecedents of goal ambiguity than those of the former stream, and some studies suggest more complicated paths to consequences. For example, Wright and Davis (2003) showed a path from organizational goal specificity to job satisfaction through human resource development; Stazyk and Goerdel (2011) reported that goal ambiguity mediates the relationship between political support and organizational effectiveness and that hierarchical authority mitigates the negative effect of goal ambiguity on effectiveness.

Studies on Program or Policy Goal Ambiguity in Public Management

By extending the unit of analysis to the program level, this study is closely connected with the first stream of empirical studies about the negative effect of goal ambiguity on performance. Even at the program level, it is assumed that goal ambiguity harms performance and that a higher level of goal ambiguity...
exists in government agencies than in private firms, for the same reasons (e.g., lack of profit indicators, more political intervention, and conflicting goals) as at the organizational level (e.g., Heinrich 1999; Matland 1995; O’Toole 1986; Ripley and Franklin 1982). In particular, Ripley and Franklin (1982) contended that goal ambiguity is one of the five salient factors affecting program performance in the U.S. federal government. The reason suggested by Matland is that “goal ambiguity is seen as leading to misunderstanding and uncertainty and therefore often is culpable in implementation failure” (1995, 158–59). Further, in a thorough literature review covering more than 100 program implementation research studies, O’Toole (1986) distinguished goal ambiguity as one of the critical antecedents of successful public programs and policies.

However, research that empirically tests the relationship between program goal ambiguity and performance has been rare. The main reason is that the difficulties of developing measures of goal ambiguity and finding performance data that are objective and comparable across a range of different types of programs have made it daunting to investigate the relationship. Accordingly, previous studies on the adverse effect of goal ambiguity on performance at the program or policy level have been almost entirely anecdotal, as some scholars have pointed out (e.g., Goggin et al. 1990; O’Toole 1986). Case studies or small-sample studies have been criticized for low external validity. The suggestions made by these scholars (e.g., Goggin 1986; Goggin et al. 1990; O’Toole 1986) for the improvement of implementation theory can be also applied to the theory of goal ambiguity at the program level: clarifying key concepts and looking at larger samples. One recent empirical study conducted by Heinrich (2012) included the replicated evaluative goal ambiguity of Chun and Rainey as a control to explain the PART results score as program performance across programs. However, this research showed a statistically insignificant relationship between evaluative ambiguity and the results score. A possible reason could be the sampling bias. Heinrich (2012) analyzed 89 programs only in the U.S. Department of Health and Human Services (DHHS), which is the agency with the second-highest percentage of programs rated as not performing well (“ineffective” or “results not demonstrated” made up 27 percent) in the U.S. federal government; compared to other programs, overall, the DHHS programs suffer from the intensive work undertaken to construct new measures (e.g., outcome measures related to evaluative ambiguity) for assessment and from limited resources (Heinrich 2012; Radin 2006). Accordingly, given these limitations of previous studies, as well as the importance of the program as a unit of analysis in the research and practice of public management, research on program goal ambiguity such as this study is warranted for a better understanding of the topics relevant to the six hypotheses suggested here.

**Developing Concepts and Measures of Program Goal Ambiguity**

Against these limitations from the previous studies, and based on Jung’s (2012a) research that developed new concepts and objective measures (explained later) of program goal ambiguity, this study identifies and measures the multiple objective dimensions of goal ambiguity with a large sample of programs from many U.S. federal agencies and analyzes how these goal ambiguity dimensions relate to a series of PART performance scores (i.e., design, planning, management, and results) of these programs. Jung defined program goal ambiguity as “the extent to which a set of goals of the federal programs allows lack of clarity in deciding work related to target, time limit, and external evaluation” (2012a, 679–80). Program or policy refers to “the governmental activities formulated in response to an authoritative decision” (Matland 1995, 154).

This definition of program goal ambiguity is simultaneously based on goal-setting theory and the PART goal system. Goal-setting theory provides the three goal components: a goal is “the object or aim of an action to attain a specific standard (target) within a specified time limit” (Latham 2004, 126); *external evaluation* is a critical part of a goal because it leads to an increase in external interactions and distractions and a decrease in work motivation and performance (Oldham and Brass 1979; Shalley 1995). These goal components were matched with targets, timeline (long-term versus annual), and evaluation (output versus outcome measures) in the PART goal system.

The three concepts and measures of goal ambiguity in this study were distinct and newly measured for use with the PART data at the program level rather than the agency level. This article briefly discusses the reasons for developing multiple measures of program goal ambiguity, comparing the measures to those of Chun and Rainey’s goal ambiguity research.

Chun and Rainey developed goal ambiguity concepts and measures using the goal statements provided by federal agencies in their strategic plans, which were required by the Government Performance and Results Act (GPRA) of 1993. Federal agencies prepared their GPRA strategic plans according to common guidelines. This provided greater uniformity of format, which made the plans comparable. PART also required a similar and comparable format of performance goals and measures among programs. PART, however, focused on federal programs, as opposed to agencies, and the information provided differed from that provided by the GPRA strategic plans. Therefore, it was necessary to develop new concepts and measures of goal ambiguity from the PART data.

The OMB (2006) guidelines for PART encouraged program representatives to state both output and outcome objectives and emphasized the importance of stating outcome objectives for the program. The guidelines encouraged the statement of targets to accompany the objectives, indicating the level of accomplishment against the objective. The guidelines also called for clear statements of both long-term and annual objectives. In developing concepts of program goal ambiguity, it was assumed that program representatives would generally seek to comply with these OMB guidelines. Chun and Rainey (2005b) implicitly assumed that if the agencies could state clear and unambiguous goals in their strategic plans, they generally would, for various reasons including the avoidance of criticism from oversight officials and other critics (refer to Mercatus Center [2007] at George Mason University).
On the basis of this commonly implicit assumption related to rational management or rationalizing reforms, but because of the different data sources, Jung (2012a) developed two new dimensions of program goal ambiguity (target ambiguity and timeline ambiguity) and replicated Chun and Rainey’s (2005b) evaluative goal ambiguity at the agency level with a slightly different name, program-evaluation ambiguity, as explained in the next section. Accordingly, this study links these dimensions of goal ambiguity to performance scores at the program level in order to contribute to the improvement of the theory of goal ambiguity.

Program Goal Ambiguity and Performance

Target Ambiguity and Performance

Target ambiguity refers to the extent to which a federal program does not specify quantitative targets for the stated goals (Jung 2013b).6 Goal-setting theorists posit that goals with clear targets lead to increased work motivation and performance (Locke and Latham 2002). The reason is that clear targets provide clear signs of accomplishment and designate the type and amount of effort required to achieve those accomplishments, which, in turn, facilitates management and enhances performance (Bandura 1989). Some public management scholars also contend that setting clear targets for performance goals is a crucial motivator and guide for public managers who provide a sense of direction to employees as they work toward achieving the performance goals or implementing policies (e.g., Boyne and Chen 2007; Heinrich 2003; Khademian 1995). In contrast, a lack of clear targets for goals indicates possible performance weakness, according to rational management. For example, once performance goals are set in the process of government reform under rational management, it is likely that the activities for the goals will be stressed and supported regardless of whether clear targets are stated; however, resources are likely to be used more inefficiently and the overall performance of the public program or the agency activities is likely to be impeded in the absence of clear targets (Hyndman and Eden 2000).

Hypothesis 1: Target ambiguity is negatively associated with program performance.

PART assumed that performance goals of individual programs needed to be specified quantitatively (OMB 2006). For some programs, the PART reports stated concrete targets for some of the goals, which included objective measures (e.g., for the Hazardous Materials Transportation Safety–Emergency Preparedness Grants Assessment in the Department of Transportation, a target for 2007 was that “the number of serious hazardous materials incidents should be no more than 466”) or, in a few cases, subjective measures (e.g., a target of overall satisfaction rate as of 2007 was that “63 percent of the respondents should assess the services from the Veterans Disability Compensation program as excellent”). The PART reports for a number of programs, however, provided no such targets for some of their goals; they stated just “N/A,” “targets under development,” “design system,” or “establish target.” Thus, target ambiguity was measured by the proportion of performance objectives without clear targets among the total number of performance objectives stated for an individual federal program.7

Timeline Ambiguity and Performance

Timeline ambiguity refers to the lack of clarity in distinguishing between annual goals and long-term goals (Jung 2012a). According to goal-setting theory, proximate goals, such as annual goals, help activate self-influence, which facilitates individual and organizational competence, whereas distant goals, such as long-term goals, provide less effective incentives to facilitate current management and performance (Bandura 1989). However, where the progressive steps to achieve long-term goals are clear, present achievements can mobilize effective management (Bandura 1989). Hence, when goals are not clearly set as long term or annual, or when long-term goals do not include clear progressive steps for their achievement (in cases of high timeline ambiguity), goal setting is less likely to be helpful for performance. In addition to clear or quantified targets, one of the important themes of rational management is that strategy should eventuate from the establishment of systematized forms of planning or time-bounded goals (e.g., long term or annual) (Hyndman and Eden 2000, 2001). Setting performance goals with clear timelines and targets is recognized by program or organizational leaders as very valuable in communicating to individuals, enhancing employee commitment to organizational goals, and improving the focus of individual teams within the program or the agency, all of which are perceived as contributing to the improvement of the management and performance of the program or agency (Hyndman and Eden 2001).8

Hypothesis 2: Timeline ambiguity is negatively associated with program performance.

In terms of time span, PART provided two kinds of performance indicators: long term (e.g., a long-term goal in the Agricultural Credit Insurance Fund Direct Loans program was “increased percentage of farm ownership by racial and ethnic minorities and women farmers should be $55 billion by FY 2007”) and annual (e.g., as an annual goal for 2007, “the volume of commodity graded through the Agricultural Grading and Certification program, measured in pounds [billion lbs], should be 89 lbs”). In addition, many programs had some performance objectives that were stated as simultaneously long term and annual. Some of these objectives presented clear explanations or progressive steps to achieve the final targets. For example, in the Disaster Response program of the Federal Emergency Management Agency, although the performance measure “casualty treatment, evacuation and transportation capability of the national NDMS [National Disaster Medical System]” was stated as both long term and annual, clear annual progressive steps were provided (e.g., 5,000 patients in 2006 and 20,000 in 2007) to achieve the long-term target (100,000 patients by 2011) (Jung 2012a). By contrast, other objectives stated as both long term and annual provided neither annual progressive steps toward the long-term target nor explanations. There may be multiple reasons for this. First, it may have to do with the difficulty of developing long-term goals—for example, the Burial Benefits program of the Department of Veterans Affairs, which represents a service to however many individuals qualify and the costs of which were driven by the mortality of eligible veterans (Moynihan 2006). Second, it may be that observing program outcomes within an annual time frame is practically impossible—specifically, in the Farmland Protection Program of the National Resource Conservation Service, the results were not complete at the time of the PART analysis, as it was still challenging to quantify annually, as well as in the long-term the environmental benefits of an agricultural easement (Moynihan 2006). Therefore, these were ambiguous goals in terms of the timeline for their...
achievement, that is, they could lead to different interpretations of the timeline for the same goal, resulting in greater goal ambiguity. Hence, the measure was calculated as the proportion of this type of performance objectives among the total number of performance objectives in individual federal programs.9

Program-Evaluation Ambiguity and Performance

Program-evaluation ambiguity is defined as the lack of clarity that a program goal “allows in evaluating the progress toward the achievement of the goal” (Chun and Rainey 2005b, 4). This concept is related to the possibility of external evaluation of performance measures or goals. It is focused on whether a program’s goal statements focus on output (or process) measures, which suggests greater ambiguity about external evaluation of performance goals, as opposed to focusing on outcomes (or results) (Chun and Rainey 2005b). That is, the logic is that the more output measures a program has, the greater its evaluation ambiguity. However, one could comment that output measures can be proximal, controllable, and clear, as compared to outcome goals; employees can more readily see the influence of their effort on output goals, as compared to outcome goals. By contrast, outcome goals can be more subject than output goals to influence from actors and conditions external to the program (e.g., politicians, interest groups, presidential salience, media salience), other than employee effort. Thus, output goals can motivate employees better than outcome goals. The problem with this interpretation is that focusing on output goals can lead to goal displacement in the classic sense, whereby one concentrates more on rules and procedural compliance and activity levels than on actual results (Meier and Bohte 2000; Merton 1940). For example, one concentrates on the number of employment trainees trained rather than whether their job prospects and their earnings improved. Considering this problem, program-evaluation ambiguity does not refer to goals that are clear in the employees’ mind but to goals that clarify to employees and external evaluators the ultimate results that the program or agency seeks to achieve.

In relation to evaluation ambiguity, the designers of PART stated that they sought greater accountability for program performance by requiring federal programs and agencies to set performance goals and measures and to make their performance results known to the public (Heinrich 2007). Thus, the provision of more outcomes and results is expected to enhance the ability of the public, legislators, and bureaucrats to evaluate program performance and the accountability of programs for their activities. One possible reason for the expected negative relationship between this ambiguity and performance is that the lack of outcomes or results for public agencies can cause leaders to pay more attention to the overall reputation of their organization and their relations with politicians and mass media than to vital performance (Allison 1983; Blumenthal 1983).10

Hypothesis 3: Program-evaluation ambiguity is negatively associated with program performance.

Although the definition and logic of this ambiguity are the same as those of Chun and Rainey’s (2005b) evaluative goal ambiguity, the measure is slightly different. The reason is that the PART reports concentrated on federal programs, as opposed to the GPRA reports’ concentration on federal agencies, and simply provided different information. That is, the information in PART simply did not support the same measure of goal ambiguity. Accordingly, program-evaluation ambiguity needed to be labeled differently. The prior evaluative goal ambiguity was measured by the percentage of workload-oriented performance indicators, as opposed to results-oriented performance indicators, in each agency (Chun and Rainey 2005b, 13). On the other hand, the PART reports included outcome and output indicators. The examples of outcome indicators are service quality (e.g., “accuracy of loan guaranty activities” in the Veterans Benefits Administration) and customer satisfaction (e.g., “percent of respondents who rate the quality of service provided by the national cemeteries as excellent” in the National Cemetery Administration). The output indicators are associated with an increase in quantity (e.g., “number of vocational rehabilitations placements with new employers” in the Employment Standards Administration) or number of participants (e.g., “number of foreign exchange participants by region to reflect current U.S. foreign policy objectives commensurate with funding” in the Department of State). Program-evaluation ambiguity was based on the point that “workload-oriented performance indicators refer to input and output indicators as opposed to such results-oriented ones as outcome and efficiency measures” (Chun and Rainey 2005b, 13). Hence, goal ambiguity was calculated by the proportion of output measures among all performance indicators for each federal program.11

Different Effects of Goal Ambiguity Dimensions on Different Performance Dimensions

The dependent variables here are the PART scores, which a host of empirical studies have used as measures of program performance (e.g., Gilmour and Lewis 2006a, 2006b; Lewis 2008). The PART performance scores included four dimensions: design, planning, management, and results. Each PART questionnaire included approximately 30 questions that were divided among the four sections (OMB 2007). (For examples of the questions, see appendix A.) The number of questions in each PART section varied according to the distinctive characteristics of each program type. The OMB (2007) explained the four sections as follows: The “design” section asks whether a program’s purpose is clear and whether it is well designed to achieve its objectives; the “planning” section involves strategic planning and weighs whether the agency establishes valid annual and long-term goals for its programs (these explanations about the design and planning scores indicate that there could be measurement issues related to the target ambiguity and timeline ambiguity measures; see appendix B); the “management” section rates the management of an agency’s program, including financial oversight and program improvement effort; the “results” section focuses on results that programs can report accurately and consistently.

Regarding the performance types of these four sections, Gilmour and Lewis (2006a) argued that the first three sections are not directly associated with program results or outcomes but gauge features of programs that might arguably lead to good results. That is, unlike the fourth section, which measures program outcomes or results, the first three mostly concern the internal management process. As some scholars have pointed out (Gilmour and Lewis 2006a; Heinrich 2012), they measure the degree to which federal programs have done a good job in producing the required paperwork or following procedures suggested by the OMB. Therefore, the design, planning, and management sections are more closely
associated with internal process models related to performance (Rainey 2009), while the results section relates more closely to the outcome of an “input-output-outcome” model of performance, which encompasses program impact and effectiveness (Boyne 2002). Given the multiple dimensions of both goal ambiguity and performance, it is expected that the dimensions of goal ambiguity can have different degrees of association with different performance dimensions. The obvious one is that, given its definition, program-evaluation ambiguity will relate more closely to the results score than to the internal management or process-oriented scores—design, planning, and management—which indicates greater focus on output rather than outcome indicators (hypothesis 4). On the other hand, target ambiguity and timeline ambiguity are expected to be more closely related to the three process-oriented performance dimensions (hypothesis 5). The reason is that creating performance standards and targets and setting short-term and long-term goals are included in the list of procedures emphasized by the GPRA (Lynn 1999). However, public management scholars have also argued or demonstrated that setting clear targets can help program members to focus on outcome rather than process and provide a sense of direction, leading to better program performance (Boyne and Chen 2007; Jung 2011). Hence, target ambiguity will be negatively related to the program results score as well as the other performance scores. However, this ambiguity is expected to have a closer association than timeline ambiguity with both the process-oriented and outcome performance dimensions (hypothesis 6). The reason is that setting clear targets is more fundamental than or occurs prior to the development of a timeline (Connell and Klem 2000).

**Hypothesis 4:** Program-evaluation ambiguity is more negatively associated with outcome scores than with process-oriented scores.

**Hypothesis 5:** Target ambiguity and timeline ambiguity are more negatively associated with process-oriented scores than with outcome scores.

**Hypothesis 6:** Target ambiguity is more negatively associated with both process-oriented and outcome scores than timeline ambiguity.

In addition to the four dimensions of performance, the current research includes the overall program assessment rating.\(^{12}\)

**Data and Measures**

**Data Sources and Sample**

The assessment scores analyzed in this article came from the PART assessment results made public by the OMB.\(^{13}\) The programs represent a variety of federal programs that differ widely and represent diverse goals and functions. The PART data supported efforts to assess program performance in order to achieve better results and implement performance budgeting (Gilmour and Lewis 2006a; Moynihan 2006; OMB 2006). The sample for this study included all of the federal programs that were assessed at least twice between fiscal years 2004 and 2008. The sample size was 188 programs. This study used information in the PART data for the dependent and some of the independent variables (e.g., three goal ambiguity dimensions, program type, assessment year, and budget).

The use of the PART data could have several advantages in terms of program performance and goal ambiguity. First, we could use the standardized terms, which reduces a frequent problem: measurement error in gathering data from several archival sources (Van de Ven and Ferry 1980). Second, it was the most acceptable way to identify goals of federal programs; all federal programs were required to prepare the data under OMB guidelines. Third, using the data allowed for cross-program comparisons in a way that few other measures have done before. Fourth, it is better than many of the frequently used alternatives, such as self-reported performance; the data also have potential biases, but controlling for biases (e.g., agency ideology and administration by commission) observed by previous studies, as explained later, can alleviate them. At least for the 2000s, PART could be one of the best workable large databases in the United States for performance across federal programs created by the OMB, which is one of the key stakeholders for federal programs, similar to local authorities’ core service performance (CSP) scores created by the Audit Commission, a key stakeholder in the United Kingdom (for details about the CSP scores, see Andrews, Boyne, and Walker 2006). The U.K. performance measures have also been shown to be associated with some biases, such as partisanship (Bertelli and John 2010), but have been a widely used performance indicator. Even with biases, chances are there is a real relationship between actual performance and the grades. Controlling for the biases increases the likelihood that any relationship found actually exists.

**Controls**

This study also took into account controls that might otherwise lead to biased estimates of the main variables.\(^{14}\) Accordingly, it controlled for program- and agency-specific factors that might affect PART scores, as described in this section. Manager discretion in program management can vary with type of program; different programs receive different levels of external intervention from diverse stakeholders, including constituencies, interest groups, and politicians (Behn 2001).\(^{15}\) Furthermore, it is worth noting that public programs find and use different policy tools for better management and performance. The capacity (e.g., knowledge and resources) to find the most appropriate tool can vary across program types (Salamon 2002). Hence, this study controlled for the seven program types suggested by the OMB: direct federal, credit, research and development, block/formula grant, competitive grant, capital assets and service acquisition, and regulatory.

In addition, the year in which the PART assessment was conducted may also have influenced the scores. The OMB (2007) explained that assessment year indicated the most recent assessment year. Program representatives could request a reassessment when they made significant changes to improve their program rating scores, which could be linked to a kind of selection bias for reassessment. Alternatively, if a particular federal program was assessed later than others, the manager of the program would have had the opportunity to develop more experience following the PART guidelines and improving PART scores to obtain larger budget allocations (Gilmour 2006). Another relevant consideration is connected to the fact that the George W. Bush administration, in press releases and budget documents, made the case that the performance of federal programs was improving by pointing to improved PART scores. Such actions could create internal and political pressure for program
representatives to provide better ratings over time, which, combined with the growing skills to present the programs favorably, could have led to higher PART scores. Because one difficulty with comparing performance across numerous programs is the dramatic difference in their size, the models controlled for each program’s actual budget; larger programs may be more challenging to manage, or a greater budget may be conducive to better management (Gallo and Lewis 2012).

Agency ideology as the policy content of what agencies do needs to be considered; in terms of mission, history, and personnel, some agencies are more liberal or conservative than others. If PART scores are politicized, the differences in agency ideology can influence program evaluation; in other words, programs in liberal agencies can receive systematically lower grades in the PART system generated by a Republican administration, if the evaluation is politicized (Gallo and Lewis 2012, 227–28). For agency ideology, this study used Clinton and Lewis’s (2008) agency preference estimates, which ranged from –1.72 (most liberal; e.g., Corporation for National and Community Service and Peace Corps) to 2.40 (most conservative; e.g., Department of the Navy).

In addition, contextual differences at the agency level that might also influence PART’s scores include structural features such as the presence of a fixed term for the agency manager and whether the agency was managed by a commission (or administration) (Gallo and Lewis 2012). These features could lead to depoliticized management and better management performance (Gilmour and Lewis 2006b, 2006c). Another structural feature of agencies that implemented federal programs was the number of programs housed in an agency; a greater number of programs in an agency could contribute to management difficulty (Gallo and Lewis 2012).

Finally, this study considered prior performance scores that might systematically influence the current performance. As described earlier, in the PART system, programs would request a reassessment in the case of making significant changes to increase performance ratings. Table 1 presents the descriptive statistics of all of the variables in this analysis.

**Method and Results**

The dependent variables data were expressed as percentages in the PART, but for analytical purposes, this study must treat them as proportions. However, proportions are not appropriate for ordinary least squares regression, for several reasons: first, proportions violate the assumptions of normal distribution and homoskedasticity because of the bounded nature of the interval; second, the conditional expectations function is not linear because of the distribution on a closed interval; third, ordinary least squares regression estimates values outside the (0,1) interval (Kieschnick and McCullough 2003; Smith and Fernandez 2010). As a solution to these problems, Papke and Wooldridge (1993) suggested a fractional logit model as a quasi-maximum likelihood estimation. This technique includes all predicted values of the dependent variable, including 0 and 1, which do not need to be adjusted prior to estimation (Smith and Fernandez 2010). Additionally, the results of the variance inflation factor test (all scores less than 3) suggested that multicollinearity did not harm the results. This was also supported by the correlation analysis, with all coefficients between independent variables smaller than .43, as shown in appendix C.

Table 2 presents the fractional logit estimation for the five dimensions of program performance. Target ambiguity had a statistically significant and negative association with PART planning, management, and results scores. The marginal effects of a 10 percent decrease in target ambiguity were increases in the scores of roughly 0.66 percent, 0.66 percent, and 1.1 percent, respectively. Timeline ambiguity and program-evaluation ambiguity each predicted just one dimension of performance—design and results, respectively—with a statistical significance. The marginal effects of a 10 percent increase in timeline ambiguity and evaluation ambiguity were decreases in the design and results scores of approximately 0.12 percent and 1.1 percent, respectively. Thus, hypotheses 1, 2, and 3 were partially supported. In terms of overall PART scores, these three dimensions of program goal ambiguity were also negatively associated, with a statistical significance. Patterns similar to these relationships between the goal ambiguity measures and performance dimensions were found in the analysis of 752 programs (results available from the author) without including each prior performance score.

Regarding different relationships between different dimensions of both goal ambiguity and performance, evaluation ambiguity was more negative for the results score than for the other process-oriented performance scores, as expected in hypothesis 4. Related to its definition, providing more output indicators than outcome indicators (higher evaluation ambiguity) could be a critical obstacle to improvement in program outcome or results. Supporting hypothesis 5, timeline ambiguity related solely to the design score. However, target ambiguity was slightly more important for the program results than for the planning and management. The results revealed that clear target and timeline settings were key factors in the list of procedures or processes emphasized by the rationalizing reforms.
Table 2: Results of Fractional Logit Estimation for Program Performance

<table>
<thead>
<tr>
<th>PART Design</th>
<th>PART Planning</th>
<th>PART Management</th>
<th>PART Results</th>
<th>Overall PART Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef.</td>
<td>z-score</td>
<td>dy/dx</td>
<td>Coef.</td>
<td>z-score</td>
</tr>
<tr>
<td>Target ambiguity</td>
<td>-0.10</td>
<td>-0.32</td>
<td>-0.09</td>
<td>-0.51*</td>
</tr>
<tr>
<td>Timeline ambiguity</td>
<td>-0.13**</td>
<td>-2.80</td>
<td>-0.012</td>
<td>-0.07</td>
</tr>
<tr>
<td>Evaluation ambiguity</td>
<td>-0.17</td>
<td>-0.48</td>
<td>-0.016</td>
<td>-0.22</td>
</tr>
<tr>
<td>Direct federal</td>
<td>-0.63</td>
<td>-1.20</td>
<td>-0.061</td>
<td>0.35</td>
</tr>
<tr>
<td>Credit</td>
<td>-0.80</td>
<td>-1.36</td>
<td>-0.094</td>
<td>0.37</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>0.40</td>
<td>0.52</td>
<td>0.031</td>
<td>0.56</td>
</tr>
<tr>
<td>Block/Formual grant</td>
<td>-0.85</td>
<td>-1.52</td>
<td>-0.097</td>
<td>0.63†</td>
</tr>
<tr>
<td>Competitive grant</td>
<td>-0.42</td>
<td>-0.67</td>
<td>-0.042</td>
<td>0.82*</td>
</tr>
<tr>
<td>Capital assets and service acquisition</td>
<td>-0.56</td>
<td>-0.97</td>
<td>-0.058</td>
<td>0.29</td>
</tr>
<tr>
<td>Assessment year</td>
<td>-0.13</td>
<td>-1.37</td>
<td>-0.011</td>
<td>0.03</td>
</tr>
<tr>
<td>Budget size</td>
<td>-0.08</td>
<td>-1.44</td>
<td>-0.007</td>
<td>-0.93</td>
</tr>
<tr>
<td>Agency ideology</td>
<td>-0.18</td>
<td>-1.38</td>
<td>-0.016</td>
<td>0.13</td>
</tr>
<tr>
<td>Administered by commission</td>
<td>0.18</td>
<td>0.64</td>
<td>0.015</td>
<td>1.03**</td>
</tr>
<tr>
<td>Fixed term for manager</td>
<td>-0.16</td>
<td>-0.28</td>
<td>-0.015</td>
<td>0.00</td>
</tr>
<tr>
<td>No. of programs</td>
<td>-0.02</td>
<td>-0.98</td>
<td>-0.002</td>
<td>0.01</td>
</tr>
<tr>
<td>Prior respective scores</td>
<td>3.99**</td>
<td>7.02</td>
<td>0.356</td>
<td>1.92**</td>
</tr>
<tr>
<td>Constant</td>
<td>0.68</td>
<td>0.80</td>
<td>0.23</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Notes: Sample size = 188; for program types, the reference category is regulatory programs. **p < .01; *p < .05; †p < .10 (two-tailed test).

(e.g., Dull 2006; Lynn 1999), but clear targets also help focus on outcomes rather than processes (Boyne and Chen 2007; Jung 2011). Then, hypothesis 6 was also partially supported. The reason is that target ambiguity was more important for planning, management, and results than timeline ambiguity, which was more important for the program design score. This could be evidence for the argument that target setting is more fundamental and important than timeline building (Connell and Klem 2000).

Turning attention briefly to the controls, they were not statistically significant in many cases, unlike the correlation results. The main reason seemed to be the inclusion of the prior performance score in each model. The analysis of program type showed that grant-based programs—block/formula and competitive grants—had higher levels of planning performance than regulatory programs. Providing evidence that the characteristics of the program, such as the tasks and policy tools involved, could be associated with performance, this analysis suggests the value of continuing research on the relationship between program type and performance. In addition, the recent assessment year related positively to the results score. This evidence could support the interpretation that there might have been a learning or development process by which program managers became better at complying with the OMB requirements and preparing PART reports about results. The presence of a commission was positively related to the externally evaluated performance scores concerning planning, management, results, and the overall rating. In addition, the negative associations of budget size and number of programs with the management score suggest the difficulties of administering larger and more complex programs and agencies. Furthermore, a fixed term for the agency manager showed an unexpected relationship with the management score. This could be evidence that the PART scores were politicized. Then, as expected, each prior performance score provided the most powerful explanation in each model. These findings could increasingly be woven together for more detailed and systematic explanations of why programs varied on performance assessment.

Discussion

The findings of this research have several implications for public administration. First, in a broad sense, consistent with prior studies, the results further support the observation, repeated time and again by prominent scholars and experts, that ambiguous goals can impede performance in government agencies. Yet, since Rainey’s (1993) theory of goal ambiguity in public agencies, the few studies of this goal characteristic have concentrated on the overall agency level, irrespective of whether they used archival data or survey data. The empirical effort on this topic has been very rare at the program level. However, because of the availability of PART data, which offered goal systems comparable across programs, this research could provide the first empirical evidence with a large sample about the negative relationship between goal ambiguity and performance at the program level.

Second, for the new evidence from the program level, this study presented three objective measures of program goal ambiguity. Target ambiguity and timeline ambiguity were newly measured at the program level, and program-evaluation ambiguity replicated Chun and Rainey’s (2005b) evaluative goal ambiguity of the agency level. Providing diverse goal ambiguity measurements is a main contribution of goal ambiguity studies using objective measures; simultaneously, it is a limitation of the other research stream, which has just measured overall organizational goal ambiguity level using survey measures. These two streams of empirical studies have made different contributions to the goal ambiguity scholarship, as discussed earlier. With these old and new dimensions of goal ambiguity, this research can contribute to the generalizability of the theory of goal ambiguity (Jung 2013b). Specifically, at the program level, this study provided evidence validating Chun and Rainey’s evaluative...
goal ambiguity measure and supported the assumption about its negative relationship with performance. Thus, researchers or theorists can now pursue more refined explanations of why certain government programs have certain levels of performance on various administrative aspects or external evaluations, as related to goal-setting activities. Specifically, setting clear targets and clear timelines and providing more outcome measures matter for improving program performance. Of the three goal-setting activities, specifying targets was the most conducive to improvement in performance.

Third, this research suggested new hypotheses (4, 5, and 6) with diverse dimensions of both goal ambiguity and performance. These hypotheses were developed to help explain the differences in expectation between what should explain the models of program internal management or process-oriented performance (i.e., design, planning, and management) and program results. That is, the importance of each goal-setting activity can vary across different administrative activities. Specifically, evaluation ambiguity related more closely to the performance evaluation of results, while the other two ambiguity dimensions are more closely associated with process-oriented performance. In addition, target ambiguity is more important than timeline ambiguity for performance evaluation.

Admittedly, this study did not hypothesize how differently the three ambiguity dimensions are associated with the four performance dimensions. Yet no goal ambiguity studies in public management have differentiated the effects of diverse ambiguity dimensions on different performance dimensions, as in this research. Therefore, testing the hypotheses and providing the expected results can lead the goal ambiguity scholarship to refine the reasons for the variations in the relationships between multiple dimensions of both goal ambiguity and performance, and how these variations occur.

With these implications, this article contributes to the development of more refined predictive models of program performance in public management. Specifically, the old and new dimensions of goal ambiguity are added as new antecedents of program performance. These results were found after considering a host of controls derived from the relevant literature. Some of the controls were associated with biases or limitations in the PART as a performance instrument. For example, PART ratings were widely criticized because of PART’s ideological bias toward republicanism or conservatism, the political preferences of the president, and the assumption that programs of different types can be evaluated against the same set of standards (Dull 2006; Gallo and Lewis 2012; Gilmour and Lewis 2006a, 2006c; Moynihan 2006). Considering these suggested biases, this research controlled for agency ideology, administration by commission, fixed term of the manager, and program type.

This study also added to previous evidence that PART showed some patterns of bias (e.g., Gallo and Lewis 2012; Gilmour and Lewis 2006a; Lewis 2008). Yet if the PART scores were seriously flawed by these problems in rating and uneven resources for assessment, arguably, the performance scores should not be associated with the goal ambiguity measures. However, these relationships held in the analysis when these factors were taken into account that could represent alternative, biasing influences on the PART scores. That is, the inclusion of these factors could enhance the credibility of the main findings here. In this way, the theory of goal ambiguity and the model of performance in public management can be advanced by this program-level research. As some authors have stated, public administration as a discipline of social science or the theory of goal ambiguity can thrive when important and interesting questions lead to effective data gathering and consequential findings (e.g., Locke and Latham 2009; Mischel 2009).

Further, there are some practical implications for government agencies. Practical and policy decisions can take into consideration variations among programs and agencies that can influence their goal ambiguity and performance scores and consider whether those variations can be taken into account in program design. Alternatively, they may be taken into account in the design of performance evaluations, which, in some cases, may need to acknowledge that some programs and activities should not seek premature specification of their goals (Gilmour 2006). Other cases can be identified in which goal clarification is feasible but underdeveloped. In this regard, critics of PART might point to an alternative explanation contending that stating clear goals and plans was simply easier for some federal programs than for others and that PART was simply biased toward awarding higher scores to programs for which goal clarification was easier. Alternatively, program managers might simply improve at developing convincing PART reports, including reports with seemingly clearer goals (related to this issue, please see appendix D).

The findings about goal ambiguity and the controls discussed here could be cited as evidence of such biases. This is not a strong alternative, however, as more clearly stated goals and targets also made it easier to detect when goals were not met, and the findings for the controls were not particularly strong and consistent as support for such an alternative explanation. However, even if PART did contain a bias toward programs where goals could be stated clearly, improved understanding of how to define and measure goal ambiguity, of factors that increased or decreased it, and of its relation to program performance could help avoid and remove such biases from assessment procedures. This study might also call for more research on the administrative craft of designing persuasive goal and administrative performance systems. By no means does this study interpret the evidence here as indicating that program goal clarification offers a panacea for boosting performance. Rather, the findings support the conclusion that appropriate effort to clarify goals can benefit at least some performance aspects (Latham, Borgogni, and Petitta 2008; Wright 2004). More importantly, the evidence here extends the stream of research seeking to clarify the concept of goal ambiguity, find ways of measuring it, and deepen our understanding of it.

Conclusion
This analysis of 188 federal programs, the prior performance assessment records of which could be tracked from 2004 to 2008
in PART, extends the stream of research (e.g., Chun and Rainey 2005b; Jung 2011) that has developed measures of goal ambiguity, including multiple dimensions, and has consistently supported the adverse effect of goal ambiguity on performance at the overall agency level. Although this research controlled for previous performance scores, and this setting should absorb any potential simultaneous effect, it should be acknowledged that simply controlling for prior performance scores does not address the impact that performance might have on subsequent goal setting. Thus, a new area for research would be to test the causal relationship between goal ambiguity and performance by collecting data for multiple time points.

Although PART was discontinued by the first Barack Obama administration, and it is uncertain whether the present study will be extended and replicated, the success reported here in developing measures of goal ambiguity at the program level, using methods similar to but newly adapted from previous studies at the agency level, justifies optimism about using similar methods with other forms of goal setting, performance measurement, and performance in government. Even in view of the uncertainties, the present analysis provides valuable evidence about PART, performance measurement and performance systems in government, and goal ambiguity (or clarity) of government activities; it contributes to the stream of research that is advancing the understanding of those topics and the influences on them.

Acknowledgments
I am grateful to Hal G. Rainey for his profound insights concerning challenging issues of this research and to Sangyub Ryu for his assistance in the validity tests of goal ambiguity measures. This research was supported by a Start-Up grant (Ref. no. 7200336) funded by City University of Hong Kong.

Notes
1. The survey items are “This organization’s mission is clear to almost everyone who works here” (reversed), “It is easy to explain the goals of this organization to outsiders” (reversed), and “This organization has clearly defined goals” (reversed).
2. Using the aforementioned three survey items to measure perceived organizational goal ambiguity in the prior studies, Wright and Davis (2003) named the measure “organizational goal specificity.” As such, in public management, scholars have used “organizational goal clarity” and “goal specificity” interchangeably as the opposite poles of goal ambiguity.
3. The possibility of potential benefits implies the nonlinear relationship between goal ambiguity and performance. Therefore, the nonlinearity was tested with the data. However, it was not found.
4. As an anonymous reviewer suggested, it is an interesting research topic to investigate how organizational goal ambiguity influences organizational performance when carried through by considering these job attitudes to be “third” variables.
5. In the stream using objective measures, very recently, Jung (2013a) found that target ambiguity and timeline ambiguity mediate the association between budget size and organizational effectiveness.
6. An anonymous reviewer raised a question about the difference between this target ambiguity and the evaluative goal ambiguity of Chun and Rainey (2005b) (or program-evaluation ambiguity of this research). These two constructs are different conceptually and methodologically. Conceptually, evaluative goal ambiguity is based on whether an agency’s goal statements focus on output and process measures, which suggests more ambiguity about external evaluation of performance goals, as opposed to focusing on outcomes and results (Chun and Rainey 2005b). However, target ambiguity is based on whether a program’s goal statements provide a clear target. Reflecting the difference in the definitions, these two constructs were measured with different coding schemes, as explained in this section. Further, as suggested in hypothesis 4, evaluative goal ambiguity or its equivalent, program-evaluation ambiguity, is expected to relate more closely to the PART results score than target ambiguity.
7. The source of the concurrent validity was the 9th Annual Performance Report Scorecard of the Mercatus Center at George Mason University as of 2007. This scorecard contained an assessment of the quality of the information in the three areas—transparency, public benefits, and leadership—in performance reports provided by 24 U.S. federal agencies to which the Chief Financial Officer Act applied (Mercatus Center 2007). The alternative measure for the concurrent validity was the “transparency” score, which consisted of four components—accessibility, readability and understandability, verifiability, and provision of performance measures (Mercatus Center 2007). Given these components and the definition of target ambiguity, a negative correlation was expected between the transparency score and the target ambiguity measure. Thus, the correlation result \( r = -.35, p < .05 \) could be evidence of the concurrent validity.
8. These arguments about the importance of setting clear targets and timeline were from the experience of the U.K. public sector, the Next Steps Initiative. Unlike Hyndman and Eden’s arguments, Lynn (1999) and Gilmour and Lewis (2006a) criticized the GPRA of the United States, which was an effort to manage for program results and performance budgeting, as proliferating paperwork by requiring agencies to create performance standards and targets and long-term and annual goals.
9. Face validity is defined as the extent to which evaluators judge that the measure is appropriate for the targeted concept on its face value (Haynes, Richard, and Kubany 1995). Concerning the validity of timeline ambiguity, it is obvious that the more goals in a program are stated as both long term and annual without any explanation or annual progressive steps, the higher the level of ambiguity in the program. Thus, the timeline ambiguity measure can be approved as having, at a minimum, face validity.
10. There could be criticism that the OMB examiners were biased toward outcome statements and gave better ratings to programs that stated more outcomes in the PART reports, as the OMB (2006) encouraged the statement of outcome measures. Thus, this study does not ignore the possibility that the result regarding the negative relationship between evaluation ambiguity and performance could reflect such bias against the performance evaluation.
11. For the convergent validity, this study used the same source—the 9th Annual Performance Report Scorecard (Mercatus Center 2007)—as for target ambiguity. The alternative measure was the “public benefits” score of the 24 U.S. federal agencies in the scorecard. This score was combined with four assessment sections—outcome-oriented goals, outcome measures, agency-affected outcomes, and linkage to costs related to public benefits—associated with performance measures (Mercatus Center 2007). From these sections, this score is expected to relate negatively to the program-evaluation ambiguity measure. The correlation result \( r = -.37, p < .05 \) showed the convergent validity of this ambiguity.
12. For the overall assessment rating score, the OMB provided the total weighted score, based on the weights of four assessment sections for PART: program design (20 percent), planning (10 percent), management (20 percent), and results (50 percent) (OMB 2007).
13. PART, provided by the OMB, is now well known and has been used often for program performance measures in the public management literature. In developing PART early in 2002, the OMB organized the Performance Evaluation Team, which made public the first draft of the questionnaire to seek public comment. Then, after reviews by independent groups (e.g., the Performance Measurement Advisory Council and the President’s Council on Integrity and Efficiency), a congressional hearing, and a workshop convened by the National Academy of Public Administration, the first version of PART was approved by the President’s Management Council on July 10 and published for use on July 16, 2002. After
rating 234 federal programs for fiscal year 2004, the OMB intended to raise
the number of reviewed programs by nearly 250 every year (Gilmour and Lewis
2006a; OMB 2004). Hence, the number of reviewed programs for the 2007
PART of this study was 794, and since 2008, the number has been nearly 1,000.
PART was discontinued after 2010 by the Obama administration.
14. Although a host of controls were considered here, it needs to be acknowledged
that the models did not reflect the complex decision-making process concern-
ing PART. The PART scores could represent data that had to fit into a very
complex decision-making process in which there were different and conflicting
views about the OMB circulars for PART. This was beyond the bounds of this
research.
15. For example, grant- and contract-based programs, such as competitive grants and
block/formula grants, experience more intervention in the policy process from
multiple stakeholders (Frederickson and Frederickson 2006).
16. Regarding the relationship between budgeting and PART scores, there have been
criticisms of the mismatch between executive analysis and congressional needs,
a large disparity between the president's budget requests and actual budgets, and
the lack of interest of Congress in the PART process (Radin 2008; White 2012).
17. The equation is as follows: E[y|x] = G(xβ) = exp(xβ)/[1 + exp(xβ)], where all
predicted values lie within the expected interval (0,1) (Smith and Fernandez 2010).

References
Alike in All Unimportant Respects? In Public Management: Public and Private
Perspectives, edited by James L. Perry and Kenneth L. Kraemer, 72–92. Palo Alto,
CA: Mayfield.
and Organizational Performance: An Empirical Analysis. Public Administration
Bandura, Albert. 1989. Self-Regulation of Motivation and Action through Internal
Standards and Goal Systems. In Goal Concepts in Personality and Social Psychology,
Brookings Institution Press.
How Performance Measures Expand Distributive Politics. Journal of Politics
72(2): 545–58.
In Public Management: Public and Private Perspectives, edited by James L. Perry
Boyne, George A. 2002. Concepts and Indicators of Local Authority Performance:
An Evaluation of the Statutory Frameworks in England and Wales. Public Money
Boyne, George A., and Alex A. Chen. 2007. Performance Targets and Public Service
Chun, Young Han, and Hal G. Rainey. 2005a. Goal Ambiguity and Organizational
Service Performance: Perspectives on Measurement and Management, edited by
George A. Boyne, Kenneth J. Meier, Laurence J. O’Toole, Jr., and Richard M.
a Theory of Change Approach to Plan Urban Education Reform. Journal of
Educational and Psychological Consultation 11(1): 95–120.
Planning and Politico-Economic Systems Resolved into Basic Social Processes.
Dull, Mathew. 2006. Why PART? The Institutional Politics of Presidential Budget
Frederickson, David G., and H. George Frederickson. 2006. Measuring the
Gallo, Nick, and David E. Lewis. 2012. The Consequences of Presidential Patronage
for Federal Agency Performance. Journal of Public Administration Research and
Theory 22(2): 219–43.
Gilmour, John B. 2006. Implementing OMB’s Program Assessment Rating Tool (PART):
Meeting the Challenges of Integrating Budget and Performance. Washington, DC:
IBM Center for the Business of Government.
Gilmour, John B., and David E. Lewis. 2006a. Assessing Performance Budgeting
at OMB: The Influence of Politics, Performance, and Program Size. Journal of
Public Administration Research and Theory 16(2): 169–86.
———. 2006b. Does Performance Budgeting Work? An Examination of the Office
———. 2006c. Political Appointees and the Competence of Federal Program
Goggin, Malcolm L., Ann O. Bowman, James L. Perry, and Laurence J. O’Toole,
Glencoe, IL: Scott, Foresman/Little, Brown.
Hassan, Shahidul, and John Rohrbaugh. 2011. The Role of Psychological Climate
on Public Sector Employees’ Organizational Commitment: An Empirical
Assessment for Three Occupational Groups. International Public Management
Validity in Psychological Assessment: A Functional Approach to Concepts and
Heinrich, Carolyn J. 1999. Do Government Bureaucrats Make Effective Use of
Performance Management Information? Journal of Public Administration Research
and Theory 9(3): 363–93.
———. 2007. False or Fitting Recognition? The Use of High Performance Bonuses
in Motivating Organizational Achievements. Journal of Policy Analysis and
Hyndman, Noel, and Ron Eden. 2000. A Study of the Coordination of Mission,
Objectives and Targets in U.K. Executive Agencies. Management Accounting
———. 2001. Rational Management, Performance Targets and Executive Agencies:
Views from Agency Chief Executives in Northern Ireland. Public Administration
Jung, Chan Su. 2011. Organizational Goal Ambiguity and Performance:
Conceptualization, Measurement, and Relationships. International Public
———. 2012a. Developing and Validating New Concepts and Measures of Program
Goal Ambiguity in the U.S. Federal Government. Administration & Society
———. 2012b. Why Are Goals Important in the Public Sector? Exploring the
Benefits of Goal Clarity for Reducing Turnover Intention. Journal of
Extending the Theory of Goal Ambiguity to Programs: Examining the Relationship between Goal Ambiguity and Performance


Appendix A. Examples of Specific Questions in the Program Assessment Rating Tool (PART) 

**Program Purpose and Design**

- Does the program address a specific and existing problem, interest, or need?
- Is the program designed so that it is not redundant or duplicative of any other Federal, State, local or private effort?
- Is the program design free of major flaws that would limit the program’s effectiveness or efficiency?

**Strategic Planning**

- Does the program have a limited number of specific long-term performance measures that focus on outcomes and meaningfully reflect the purpose of the program?

**Program Management**

- Does the program use strong financial management practices?

**Program Results/Accountability**

- Does the program (including program partners) achieve its annual performance goals?
- Does the program demonstrate improved efficiencies or cost effectiveness in achieving program goals each year?
- Does the performance of this program compare favorably to other programs, including government, private, etc., with similar purpose and goals?
- Do independent evaluations of sufficient scope and quality indicate that the program is effective and achieving results?

Appendix B. Measurement Issues between Goal Ambiguity and PART Performance Measures

One could point out the closeness or the possibility of tautology between target ambiguity and design performance measures and between timeline ambiguity and planning performance measures. Thus, one could contend that target ambiguity and timeline ambiguity are closely associated with design and planning performance scores, respectively. This could also be a type of common source bias; however, the possibility is not high, for several reasons. First, goal ambiguity and performance are disparate concepts, as explained in the early sections of this article. Second, the goal ambiguity measures and performance scores were gauged by different people and in different ways: The former were evaluated by the author by focusing on the proportion of ambiguous goals to the total number of goals in each program but the latter by the OMB evaluators, based on several standardized questions whose answers were coded just “yes” or “no,” like dummy coding. Third, test results provide statistical evidence about the discriminant validity between target ambiguity, timeline ambiguity, design score, and planning score. Chi-square difference tests were conducted between these measures for any possible combinations by using structural equation modeling by the M-Plus program. These results suggest that all the dimensions were distinct with a highly statistical significance: All p-values were lower than .001. These two goal ambiguity measures were also distinct from management and results scores with a highly statistical significance from the same tests. Moreover, if they were not distinct and there were some biases, the target ambiguity and timeline ambiguity measures should have the strongest association with design and planning, respectively. Yet the target ambiguity measure related more closely to the results score than design score. In addition, timeline ambiguity was not statistically associated with the planning score but solely with the design score.
Appendix C. Correlations of All Variables

|          | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1. Overall score | .93 |
| 2. Results score | .55 .34 |
| 3. Management score | .63 .47 .38 |
| 4. Planning score | .53 .29 .19 .38 |
| 5. Design score | –.37 –.35 –.22 –.25 –.15 |
| 6. Target ambiguity | –.22 –.18 –.10 –.16 –.21 .17 |
| 7. Timeline ambiguity | –.22 –.22 –.08 –.11 –.09 .09 –.04 |
| 8. Evaluation ambiguity | –.13 –.09 –.14 –.03 –.11 –.14 .09 .08 |
| 9. Direct federal | .07 .07 .14 –.02 –.07 .02 –.09 –.08 –.20 |
| 10. Credit | .18 .17 .03 .10 .16 –.07 –.06 .08 –.19 –.07 |
| 11. R&D | –.04 –.01 –.03 –.05 –.09 .04 .00 –.09 –.29 –.11 –.11 |
| 12. Block grant | –.05 –.11 .03 .03 .07 .25 .16 –.14 –.33 –.12 –.12 –.19 |
| 13. Competitive grant | .04 .06 –.01 .02 .00 .03 –.14 .10 –.29 –.11 –.11 –.17 –.19 |
| 14. Capital assets | .08 .04 .10 –.02 .14 –.15 –.06 .03 –.20 –.07 –.07 –.11 –.12 –.11 |
| 15. Regulatory | –.18 –.19 –.02 .13 .10 .01 .06 .03 –.04 –.12 .05 .04 –.07 |
| 16. Assessment year | .04 .12 –.13 –.01 –.12 –.07 –.12 –.01 .02 .06 –.02 .16 –.28 .20 –.16 –.15 |
| 17. Budget size | .05 .08 .02 .05 .07 .10 –.09 .24 .25 .12 .02 –.33 –.34 .21 .07 .01 .02 |
| 18. Ideology | .21 .22 .18 .16 –.02 –.14 .03 –.02 .09 .09 –.02 –.04 –.15 .11 .05 .12 .01 .02 |
| 19. Commission admin. | –.09 –.11 .11 –.07 –.09 .25 .16 –.09 –.28 .11 –.02 .13 .33 –.04 –.20 .04 .06 –.32 –.17 –.10 |
| 21. No. of programs | .49 .55 .11 .27 .11 –.41 –.19 –.14 .04 .07 .27 –.12 –.29 .03 .14 –.19 .19 .22 .11 .13 –.29 .93 |
| 22. Prior overall results | .26 .24 .33 .11 .00 –.23 –.06 –.03 .08 .24 .13 –.14 –.16 –.04 –.04 .07 –.02 .19 .04 –.01 –.06 .53 .32 |
| 23. Prior management | .42 .37 .16 .38 .28 –.26 –.18 –.12 –.01 –.05 .24 –.16 –.09 .07 .08 –.18 .16 .06 .06 .08 –.17 .65 .54 .20 |
| 24. Prior planning | .35 .25 .05 .21 .60 –.15 –.10 –.13 –.06 .04 .12 –.06 .04 –.01 .09 –.07 –.06 –.07 –.06 .15 .00 .44 .21 .14 .31 |

Note: All correlation coefficients above or equal to |.14| are significant at the .05 level.

Appendix D. “Goal Ambiguity” or “Compliance”

This study measured three types of program goal ambiguity regarding target, timeline, and external evaluation. As described earlier, these three goal contents were what the OMB guidelines encouraged program representatives to follow for the PART reports. Thus, these measures could possibly be used for operationalizing some types of rules or procedural compliance, which indicates the extent to which programs or their PART reports complied with the guidelines of the assessment agency. That is, the reversed measures could indicate the levels of compliance of the federal programs. If so, the hypothesis could be that when a program complies better with the guidelines of the assessment agency, it is more likely to have higher levels of performance evaluation scores. Accordingly, the results could reveal that in the George W. Bush administration, some federal programs could have concentrated on complying with the guidelines or rules in the PART system in order to get better performance scores rather than improve meaningful and desired results for the public. In the classic sense, this problem is also associated with goal displacement: some programs could focus more on making quantifiable performance goals or measures with much paperwork rather than setting those that were more difficult for employees to expect, provide, and achieve the targets but were perhaps more important for the public welfare. This could be related to why scholars and practitioners doubted the validity of the PART performance system and why the PART system was discontinued by the Obama administration.
Copyright of Public Administration Review is the property of Wiley-Blackwell and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.