

How the unintended consequences of organizational interventions complicate the assessment of economic utility

Kevin R. Murphy*

Lamorinda Consulting LLC. and Colorado State University, USA

Received: 25 September 2012

Revised: 28 November 2012

Accepted: 28 November 2012

Abstract

Interventions in organizations (e.g., implementing new testing, training or leadership development programs) are likely to have a wide range of effects, some intended and some unintended. These outcomes are likely to unfold over time in a wide range of trajectories. A multi-stakeholder, multivariate longitudinal perspective is suggested as a way of reflecting more broadly the range of effects of organizational interventions when estimating their financial impact.

Keywords: utility, financial returns, interventions, human resources

JEL Classification Codes: M12, M50, M51

1. Introduction

When consultants or managers suggest changes to organizational policies, procedures or working methods, it is reasonable to ask whether these changes are likely to benefit the organization financially. For example, an organization that considers adopting a new training program, a modification to their performance appraisals, a new set of selection tests, a new leadership development program etc. is likely to consider the projected benefits of these interventions and to compare them to the projected costs.

There is a long history of research on the best methods of estimating the economic utility of workplace interventions (Boudreau, 1991; Boudreau, Sturman and Judge, 1994). Building on the work of Brogden (1949), Cronbach and Gleser (1965), and Taylor and Russell (1939), methods of estimating the financial utility of organizational interventions

* E-mail: krm10@me.com.

Citation: Murphy, K.R. (2012) How the unintended consequences of organizational interventions complicate the assessment of economic utility, *Economics and Business Letters*, 1(4), 33-42.

such as personnel testing and training quickly developed (Cascio, 1993). Utility estimation methods have been widely applied in the field of personnel selection (Hunter and Hunter, 1984; Hunter and Schmidt, 1982; Martin and Raju, 1992; Schmidt and Hunter, 1999; Schmidt, Mack and Hunter, 1984). A number of studies have extended the basic utility estimation methods to deal with topics such as rejected selection offers and the use multiple outcome variables (e.g., Murphy, 1986; Murphy and Shirella, 1997; Winkler, Köenig and Kleinmann, 2010).

Personnel selection is not the only area in which estimated of the financial impact of organizational interventions have been applied. Methods for evaluating the financial impact of productivity enhancement programs and of changes in pay policy have been examined by Klaas and McClendon, (1996), Pritchard (1990) and Roth (1994). Boudreau and Ramstad (2003) have extended this work to assessments of strategic human resource management.

In addition to studies of methods for estimating the economic impact of organizational interventions, there has been a lively literature dealing with the credibility and meaning of these estimates to users (e.g., Bridgeman, Burton and Cline, 2009; Hazer and Highhouse, 1997; Whyte and Latham, 1997). One conclusion from this literature is that simply presenting financial projections to end users without a clear and compelling description of how interventions in organizations actually deliver these benefits tends to undermine the credibility of projections of financial benefits. A second conclusion from this literature is that the many widely used methods of estimating the financial utility of organizational interventions lack credibility because they fail to include the full range of outcomes that can be reasonably be expected when changes are made in organizational policies, practices or resources.

As Murphy and Shirella (1993) noted, most studies of the impact of organizational interventions adopt a univariate focus, choosing one variable (e.g., projected increases in productivity) to represent the principal outcome of that intervention. Organizational interventions, however, almost always have an impact on a number of outcomes, and these different outcomes will not always have consistent effects on an organization's bottom line. For example, the decision to use cognitive ability tests¹ in personnel selection is likely to lead to higher levels of productivity, but also to higher levels of vulnerability to charges of employment discrimination (Murphy, 2009). The multivariate validation framework developed by Murphy and Shirella (1993), coupled with applications of multiattribute evaluation methods (Edwards and Newman, 1982), provides an avenue for evaluating the overall effect of an intervention that can be expected to impact multiple outcome variables, but the application of these methods assumes that the full set of effects of any organizational intervention can be specified and measured.

In this paper I note that virtually any organizational intervention is likely to have a range of unintended consequences, and that the major challenge in estimating the economic impact of these interventions lies in describing as fully as possible the consequences these interventions are likely to have. I will illustrate the range of consequences (intended and unintended) that might need to be considered in evaluating the impact of an intervention. I will use as an example the most frequent, and seemingly

¹ Tests of general or specific aspects of intelligence are often referred to as cognitive ability tests. The Wonderlic Cognitive Ability Test is an exemplar of this class of tests.

most simple example of traditional utility analyses in organizations – i.e., the decision to use a valid test for selecting among job applicants. This example, then serves as a springboard for articulating a general approach toward identifying and quantifying the unintended consequences of interventions in organizations.

2. Intended and unintended consequences of testing

One reason that so much of the on the financial consequences if organizational interventions uses simple univariate methods is that the effects of interventions seem straightforward. Well-developed tests of cognitive ability yield scores that are positively correlated with measures of job performance and performance in training in virtually all jobs (Schmidt and Hunter, 1999). As a result, the use of these tests to select among job applicants is very likely to lead to the selection of job applicants with the potential to be better performers than would be selected using other less valid methods. This is also a domain in which at least one unintended effect of using this particular category of tests in selection is well understood. The use of cognitive ability tests in personnel selection is very likely to lead to accusations of race discrimination, particularly if the number of applicants is large relative to the number of people selected (Murphy, 2009). Increased productivity and increased vulnerability of employment discrimination litigation are the most frequently discussed outcomes of the decision to introduce a cognitive test into an organization's selection system, but they do not exhaust the set of outcomes that is likely to influence the financial impact of testing.

In order to understand both the intended and unintended consequences of an organization intervention (such as the decision to use ability tests in selection), it is useful to adopt a multi-stakeholder perspective (Edwards and Newman, 1982). The decision on the part of an organization to use cognitive tests in its selection program has implications for the organization, for the applicants, and for the broader society in which the organization is embedded.

Applicant perspectives. The methods organizations use to attract and screen job applicants can influence their perceptions of the organization, and perhaps their willingness to apply for jobs and accept job offers. There is evidence that the use of selection methods that are not perceived by applicants as job-related can reduce their interest in applying to and accepting job offers from organizations (Hausknecht, Day and Thomas, 2004). On the other hand, if the tests used by organizations are perceived as both relevant and fair, they might contribute to the overall perception of the organization as a desirable place to work.

Once selection decisions have been made and job offers have been accepted, the procedures used to select among applicants could still have residual effects. The processes used to attract and select among applicants have the potential to create a powerful first impression about the organization, and new employees whose initial impression of the organization is unfavorable may be less likely to develop into committed employees who perform to their potential.

The impact of selection methods on applicants' perceptions of organizations could have either long-lasting or transient effects, depending largely on the strength of the initial impression that is created and the consistency or the inconsistency between the

initial experience of testing and subsequent experiences in the organization. For example, suppose an organization used a very abstract measure of general cognitive ability as part of their selection program. Even though the empirical validity of this type of test might be acceptable (Jensen, 1980), its lack of apparent job-relatedness could create a negative impression of the organization. If an employee's subsequent interactions with the Human Resources department of his or her organization reinforce the perception that the organization does not care about or respect employees, this initial impression could have a long-lasting effect. One implication in evaluating the likely economic impact of interventions in organizations is that it will be necessary to estimate both the initial potency of the effects of the intervention and the length of time over which these effects are likely to have an influence on the behavior of job incumbents.

Societal effects of selection. The decisions organizations make about who to hire and who not to hire can have broad societal implications, particularly in context where similar tests are used in a number of settings. For example, cognitive tests are widely used for selection into college, law school, medical school and the like (Gottfredson, 1986). Because the average scores received by White examinees are typically higher than those received by African-American or Hispanic examinees, the use of these tests in selection can create a significant barrier to minority candidates. That is, if employers consistently use this class of tests to make decisions about candidates, the net effect could be a substantial under-representation of minority group members in learned professions or in top jobs. Even if each organization that uses these tests benefits from using valid tests to select from among applicants (Schmidt and Hunter, 1999), it can be argued that there are broad societal costs from this selection policy.

Putting aside the general societal costs of occupational segregation that can occur as a result of the widespread use of these tests, organizations can incur substantial costs if their employment practices are challenged as racially discriminatory. The law allows organizations to use valid, job-related tests even if they lead to differences in the likelihood of selection across protected groups (e.g., differences as a function of race, sex, age, etc.), but the process of defending even a well-validated test is a stringent and costly one. Even if selection systems are designed and intended to be race-neutral, the fact that these selection systems can have adverse impact on the employment opportunities of specific subgroups in the population presents at least the risk of substantial costs (both monetary costs and damage to the organization's reputation) if they are challenged as racially discriminatory.

Workgroup perspectives. The use of valid selection tests should lead to hiring individuals who perform the job better, learn more readily, and are more prepared to advance in organizations. In theory, these characteristics should all benefit the organization, but their effects on the immediate workgroup, and in particular on the relationships between supervisors and subordinates in the workgroup, might not be so simple. Imagine that the organization hires a cadre of true stars. This will certainly yield benefits in the beginning, but over time, these stars are likely to expect more responsibility and more advancement, and their relationships with co-workers (especially those who are more experienced but not so talented) could become increasingly strained. Similarly, supervisors might find highly talented new employees easier to lead at first, but increasingly difficult to lead at

some later point, especially if these employees come to believe that they no longer need supervision.

Time is likely to be a very important variable in evaluating the effects of using valid selection tests to add staff to existing work groups. The effects of valid testing are likely to be strongly positive at first, but they have the potential to evolve in unpredictable ways, especially in an organization that does not provide the sort of room for growth and advancement that highly talented employees might expect.

3. A general approach for modeling unintended consequences

In general, any intervention in an organization is likely to have multiple outcomes, some easily predictable (especially those that are consistent with the intended purpose of the intervention) and others both unintended and difficult to predict. The various effects of organizational interventions might unfold in quite different trajectories over time. Some intervention effects might decay over time, but others might be reasonably permanent or even might increase over time. This suggests the need for approaches that are both multivariate and longitudinal.

A general approach to identifying and taking into account the unintended consequences of organizational interventions when evaluating their overall economic utility requires the specification of five terms: (1) the outcomes – what variables are likely to be affected by this intervention, (2) the values – how do changes in these criteria map onto economic utility, (3) redundancy – to what extent are changes in outcome A as the result of an intervention consistent with changes in outcome B, (4) confidence - how confident can you be that an intervention will lead to a particular outcome, and (6) time frame.

Outcomes. Most interventions in organizations will have multiple outcomes. For example a new performance appraisal system might be introduced with the intention of providing better feedback (this intention might or might not be accomplished), but if it replaces an old one that was seen as users as reasonably satisfactory, it can also lead to more negative attitudes toward performance appraisal. There is no sure method for identifying all of the unintended consequences of organizational interventions, but a good starting point is a stakeholder analysis.

Stakeholders are individuals and groups who have a valid interest in the outcome of an intervention, and whose evaluation of whether that outcome is consistent with their own interests could reasonably affect the financial utility of that intervention. Most organizational interventions will have a set of stakeholders who include the workers directly affected by the intervention, co-workers whose own performance depends on the job performance of those directly affected by the intervention, management, the shareholders or owners of the organization, and the community that surrounds and supports the organization. By thinking through who has a valid interest in the outcome of an intervention, and what the nature of their interest is (or might be) it is often possible to identify a range of outcomes that could influence the economic utility of an intervention.

For example, the decision of whether or not to adopt an early identification program for potential organizational leaders could influence the interests of each of the stakeholder groups noted above in a variety of ways. Some of the possibilities are noted

in Table 1. These do not necessarily exhaust the domain of outcomes that might influence the economic utility of this program, but by taking the perspectives of multiple stakeholders, it is often easy to identify a set of considerations that can all influence economic utility.

Table 1. Outcomes relevant to different stakeholders if early identification program is adopted

Stakeholder	Interest	Implied Outcome Variables
Workers involved	Fair opportunity for advancement	Perceived fairness and accuracy
Co-workers	Minimize disruption	Sustained ability to accomplish tasks
Management	Identify the best candidates, Loss of best subordinates	Perceived fairness and accuracy, potential loss of high performers if they are moved into other developmental opportunities
Community	Fairness to all groups	Promotion equity across groups
Shareholders	Increased organizational effectiveness	Effective succession

Values. Mapping the outcomes that are of valid interest to each of the stakeholders onto a financial return scale will be easier for some interventions and outcomes than for others. As noted earlier, there is an extensive literature on the validity and utility of cognitive ability tests that is designed to estimate the dollar value of the increased productivity of cognitive tests. These estimates can often be done on the basis of simple linear regression equations, and while many complexities exist in this estimation process (e.g., estimating the impact of the decisions of some applicants to decline job offers), the principle that simple linear regression can forecast the value of future productivity increases is widely accepted.

Other outcomes can be mapped onto a financial value scale by specifying the process by which these outcomes are translated into actual costs or benefits. For example, if the intervention is likely to be seen by some members of the community as producing unfair outcomes, there is some risk of litigation. Exact costs are likely to be hard to nail down but it can often be feasible to estimate the minimum and likely maximum cost if litigation is indeed threatened. The minimum is given by totaling the costs of assembling and evaluating the information legal departments are likely to need to respond to litigation, whereas maximum costs are likely to be a multiple of the actual maximum damages that could be suffered by aggravated parties, plus the costs of mounting a defense. Because the possibility that: (1) litigation will ever occur, (2) that it will go to trial, and (3) that the organization will lose, are all less than 1.0 it will be necessary to accept a good deal of uncertainty in these estimates, and to model a range of best and worst-case scenarios.

Multi-attribute evaluation methods are likely to prove useful for mapping outcomes that are not in themselves financial onto a value scale that is monotonic with dollars (Edwards and Newman, 1982; Murphy, 2009). Similar to the method suggested above

for estimating the financial risks associated with litigation, these methods often begin with specification of worst and best case scenarios, with an effort to anchor those to a dollar metric. If this can be accomplished, mapping the rest of the range of plausible outcomes onto the same scale is a reasonably straightforward process.

Redundancy. It is important to identify, where possible, the likely levels of intercorrelations among the various outcomes of an intervention. There is an important distinction between a truly multivariate model (e.g., Murphy and Shiarella, 1993) and multiple independent univariate models, and if outcomes are intercorrelated (in general, they probably will be), the relationships among these outcomes will influence the evaluation of the overall effects of the intervention. In general, positive correlations among the outcomes will both simplify and amplify the effects of an intervention. That is, an intervention that has even a small positive effect on multiple, positively correlated outcomes will have a large net effect. Negative correlations among outcomes may lead to much more mixed assessments, in which the overall estimate of financial utility depends on the value associated with each outcome.

Time frame. It is important to identify the trajectory of change over time. In traditional utility models (e.g., Schmidt, Mack and Hunter, 1984), it is common to estimate the projected benefit per year, and then multiply this figure by the average tenure of employees. This is fine if one accepts the assumption that the effects of the intervention do not change over time, but this is not likely to be true for most interventions.

In general, estimates of financial utility for organizational intervention are likely to be more challenging when long time frames are considered, but the simple length of time used in estimating intervention effects is not the only relevant variable. It is likely that the effects of organizational interventions are most robust at or about the time the intervention commences, and that they decay over time. Furthermore, the rates of the decay in intervention effects are likely to vary. This suggests the need to model a range of time frames and to consider a range of decay rates when evaluating the likely financial effects of organizational interventions.

Certainty. Finally, there is plenty of room for uncertainty in estimates of the financial impact of organizational interventions, and it will be important to indicate not only a point estimate of that impact over time, but also the range of plausible values that might be observed depending on assumptions about the intercorrelations among outcomes, the trajectories of change over time, the initial size of the intervention effect, etc. Interventions will almost always have a range of intended and unintended effects that will change over time, and the best methods for estimating the financial effects of organizational interventions will take into account the multiplicity and the complexity of the effects of these interventions.

There is a large and robust meta-analytic literature that can be used to estimate the impact of a number of organizational interventions on a range of criteria. This literature is most fully developed for interventions such as testing (Schmidt and Hunter, 1999, review nearly a century of relevant research) and training (Arthur, Bennett, Edens and Bell, 2003), but well-designed meta-analytic studies of the effectiveness of interventions ranging from occupational stress reduction programs (Richardson and Rothstein, 2008) to

organizational development efforts (Neuman, Edwards and Raju, 1989) are available that provide credible data for estimating both the expected effect of interventions on specific outcome variables and the range of plausible effect sizes for those interventions.

4. Summary

Virtually any intervention in an organization will have multiple effects, some intended and some unintended. These effects are likely to unfold over time in a variety of ways, some decaying quite rapidly, some more slowly, and some trending in one direction at one point in time and in another direction at some later point in time. In order to fully capture the effects of interventions on and organization's bottom line, multiple stakeholders and multivariate longitudinal models are needed.

Virtually any point estimate of the financial effects of organizational interventions is likely to be deficient. There are too many unknowns, and estimates that incorporate error bands or that lay out the conditions under which large versus small financial effects are likely to be observed are more likely to be useful than point estimates.

References

- Arthur, W., Bennett, W., Edens, P.S. and Bell, S.T. (2003) Effectiveness of training in organizations: a meta-analysis of design and evaluation features, *Journal of Applied Psychology*, 88, 234-245.
- Boudreau, J.W. (1991) Utility analysis for decisions in human resource management, in M. Dunnette and L. Hough (eds.), *Handbook of industrial and organizational psychology*, 2nd ed., Vol. 2, 621-745, Palo Alto, CA: Consulting Psychologists Press.
- Boudreau, J.W. and Ramstad, P.M. (2003) Strategic HRM measurement in the 21st century: from justifying HR to strategic talent leadership, in M. Goldsmith, R.P. Gandossy and M.S. Efron (eds.) *HRM in the 21st century*, 79-90, New York: John Wiley.
- Boudreau, J.W., Sturman, M.C. and Judge, T.A. (1994) Utility analysis: what are the black boxes, and do they affect decisions?, in N. Anderson and P. Herriot (eds.) *Assessment and selection in organizations: Methods and practice for recruitment and appraisal*, 77-96, New York: Wiley.
- Bridgeman, B., Burton, N. and Cline, F. (2009) A note on presenting what predictive validity numbers mean, *Applied Measurement in Education*, 22, 109-119.
- Brogden, H.E. (1949) When testing pays off, *Personnel Psychology*, 2, 171-183.
- Cascio, W.F. (1993) Assessing the utility of selection decisions: theoretical and practical considerations, in N. Schmitt and W.C. Borman (eds.) *Personnel selection in organizations*, 310-340, San Francisco: Jossey Bass.
- Cronbach, L.J. and Gleser, G.C. (1965) *Psychological tests and personnel decisions*, 2nd ed., Urbana: University of Illinois Press.
- Edwards, W. and Newman, J.R. (1982) *Multiattribute evaluation*, Beverly Hills: Sage.
- Gottfredson, L. (1986) Societal consequences of the g factor in employment, *Journal of Vocational Behavior*, 29, 379-410.

- Hausknecht, J.P., Day, D.V. and Thomas, S.C. (2004) Applicant reactions to selection procedures: an updated model and meta-analysis, *Personnel Psychology*, 57, 639-683.
- Hazer, J.T. and Highhouse, S. (1997) Factors influencing managers' reactions to utility analysis: effects of SD_y method, information frame, and focal intervention, *Journal of Applied Psychology*, 82, 104-112.
- Hunter, J.E. and Hunter, R.F. (1984) The validity and utility of alternative predictors of job performance, *Psychological Bulletin*, 96, 72-98.
- Hunter, J.E. and Schmidt, F.L. (1982) Fitting people to jobs: the impact of personnel selection on national productivity, in M. Dunnette and E. Fleishman (eds.) *Human performance and productivity: human capability assessment*, Hillsdale, NJ: Erlbaum.
- Jensen, A. R. (1980) *Bias in mental testing*, New York: Free Press.
- Klaas, B.S. and McClendon, J.A. (1996) To lead, lag, or match: estimating the financial impact of pay level policies, *Personnel Psychology*, 49, 121-141.
- Martin, S.L. and Raju, N.S. (1992) Determining cutoff scores that optimize utility: a recognition of recruiting costs, *Journal of Applied Psychology*, 77, 15-23.
- Murphy, K. (1986) When your top choice turns you down: effects of rejected offers on selection test utility, *Psychological Bulletin*, 99, 133-138.
- Murphy, K. (2009) Validity, validation and values, *The Academy of Management Annals*, 3, 421-461.
- Murphy, K.R. and Shiarella, A. (1997) Implications of the multidimensional nature of job performance for the validity of selection tests: multivariate frameworks for studying test validity, *Personnel Psychology*, 50, 823-854.
- Neuman, G.A., Edwards, J. and Raju, N. (1989) Organizational development interventions: a meta-analysis of their effects on satisfaction and other variables, *Personnel Psychology*, 42, 461-489.
- Pritchard, R.D. (1990) *Measuring and improving organizational productivity*, New York: Praeger.
- Richardson, K.M. and Rothstein, H.R. (2008) Effects of occupational stress management intervention programs: a meta-analysis, *Journal of Occupational Health Psychology*, 13, 69-93.
- Roth, P.L. (1994) Multi-attribute utility analysis using the ProMES approach, *Journal of Business and Psychology*, 9, 69-80.
- Schmidt, F.L. and Hunter, J.E. (1999) The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings, *Psychological Bulletin*, 124, 262-274.
- Schmidt, F.L., Mack, M.J. and Hunter, J.E. (1984) Selection utility in the occupation of U.S. Park Ranger for three modes of test use, *Journal of Applied Psychology*, 69, 490-497.
- Taylor, H.C. and Russell, J.T. (1939) The relationship of validity coefficients to the practical effectiveness of tests in selection, *Journal of Applied Psychology*, 23, 565-578.
- Whyte, G. and Latham, G. (1997) The futility of utility analysis revisited: when even an expert fails, *Personnel Psychology*, 50, 601-610.

Winkler, S., Köenig, C. J. and Kleinmann, M. (2010) Single-attribute utility analysis may be futile, but this can't be the end of the story: causal chain analysis as an alternative, *Personnel Psychology*, 63, 1041-1065.