# THE OUTPUT CONCEPT AND PUBLIC SECTOR SERVICES

Marc Robinson,
Queensland University of Technology,
Brisbane, Australia.
Email: marc_laurence_robinson@hotmail.com or m.robinson@qut.edu.au

## Abstract

Units of output are sometimes defined in terms of the achievement of some pre-defined outcome, or alternatively in terms of the completion of some pre-defined set of activities (reflecting a *conformance to specifications* conception of quality). For most public-sector outputs, these definitions of a unit of output are flawed and may have undesirable behavioral consequences. Output measures cannot, in general, do double duty as outcome measures. Outcomes need to be measured separately. The activity content of many types of outputs may legitimately vary both over time (as a result of qualitative rationing arising from the budget constraint), and also between clients (as a result of tailoring to varying client needs). Only for a sub-set of services is it appropriate to define a unit of output as complete only when either a specified proximate outcome has been achieved, or alternatively when a pre-defined minimum set of activities has been carried out.

## Introduction

The predominant contemporary framework for conceptualizing and measuring public sector performance is the Program Logic Framework. Within this framework, the most fundamental performance concepts are inputs, activities, outputs and outcomes. The concepts of outcomes, inputs and activities are relatively clear-cut, even if the measurement challenges are sometimes considerable. There is, however, some lack of clarity both in the literature and amongst practitioners about the precise meaning of the output concept. This lack of clarity relates particularly to two inter-related matters: the distinction between outputs and outcomes and the distinction between outputs and activities. There is also a lack of understanding on the part of some concerning the distinction between outputs and what might be called *support services*.

The output concept has been important and much-used. Output measures have been a key element in the performance reporting of government organizations. There has also been a vast body of empirical work over the years by economists on public sector production and cost functions which purportedly relate outputs to inputs, and outputs to costs.

In addition, many performance budgeting systems assign a key role to outputs in the budgeting system. A central feature of the program budgeting systems which were established around the world from the 1960s was budgetary expenditure classification in terms of output-based *programs*. These were intended to facilitate effective allocative

planning of public expenditure. Contemporary performance budgeting systems generally continue to incorporate this allocative element of program budgeting. Moreover, many contemporary systems seek to go beyond program budgeting in further tightening the results-resources linkage – including the linkage between outputs and resources – by formally linking budget allocations and output and outcome performance targets. In some cases, the intention is to explicitly use output cost information (e.g. output unit costs) so as to base use agency budget funding explicitly upon planned or actual output quantity supplied (Robinson, 2002a).

Definitional issues concerning performance concepts could be regarded as matters of mere pedantic interest if their only relevance was to the labels which are attached to performance indicators. However, the wider role of the output concept in public management makes conceptual clarity no mere pedantic matter. If we are to base budgetary and other performance incentives explicitly upon outputs and outcome measures, then we need to be aware of the behavioral implications of the way in which we define these key performance concepts. Moreover, if we are to develop better *theory* on performance budgeting, clarity about the fundamental performance concepts is absolutely essential.

The primary purpose of this paper is to clarify the output concept in the service environment which characterizes the public sector. The paper also discusses some of the behavioral implications arising from the choice of output concept.

## The Output Concept and its Relation to Other Key Performance Concepts

Of the key concepts in the Program Logic Framework, only the *outcome* concept is in any sense specific to the public sector. It is not the same as the economist's concept of utility, and there are interesting unexplored issues about the relationship between the two. The outcome concept is, nevertheless, quite well-defined, even if the measurement of outcomes is often difficult: the intended outcomes of public programs are the desired changes brought about by the program upon individuals, social structures or the physical environment. A distinction is commonly made between proximate and high-level outcomes[1]. For example, student learning is a proximate outcome of school education, whereas a more productive economy is a high-level outcome.

The concepts of inputs and outputs are, by contrast, derived directly from economic analysis, in the context of which outputs usually refer to physical goods, and are synonymous with *products*. An *input*, on the other hand, is a resource (labor, materials, fixed assets) used in the production process, The further concept of *activities* – developed more by management accountants rather than economists – refers to types or categories of work task.

In the standard economic analysis pertaining to physical goods, outputs are quite distinct both from the utility (and outcomes) which they generate for individual consumers, and also from the activities which produce them. Thus, for example, no-one

would confuse a pharmaceutical (an output) with the utility or outcome (improved health) a patient derives from using it. Nor would they confuse pharmaceutical production with the pharmaceutical itself.

The concept of an output is less straightforward in the case of services than in the case of physical goods. This is of particular relevance to the public sector because most publicly-produced outputs are services rather than goods. In the case of services, there is no physical object which constitutes the output. The service provider acts directly upon a client or subject[2], with the objective of bringing about a change (proximate outcome[3]) in that client/subject. This gives the client/subject a direct involvement in the production process which does not exist for goods. These distinctive characteristics of services have led to perennial difficulties in distinguishing outputs from outcomes, and outputs from activities. As Burkhead and Miner put it in their classic work *Public Expenditure* (1971, p. 301), "the essence of the problem is that for services the distinction between the product and the effects or consequences of the product is blurred on the demand side, while the distinction between the product and the production process is likewise blurred on the supply side."

In the past, one school of analysts responded to this difficulty by using the term *outputs* to refer to what we have defined here, consistent with predominant contemporary usage, as outcomes. For example, the noted social economist Anthony Culyer (1983, p. 15) defined a service output as "the object of production … ends such as a 'healthier' patient or 'better educated' student"[4]. There are still some economists – perhaps

particularly health economists[5] – who continue to use the term output in this way. Many of those who use the term 'output' in this way continue to make use of the term 'outcomes'. Some do so without in any way distinguishing the two concepts, while others use the terms 'outputs' to refer only to proximate outcomes[6]. All of this perhaps makes it more understandable that non-experts so often confuse the output and outcome concepts – as exemplified by the former US Vice-President Gore's statement that "'outputs' are, quite simply, measures of how government programs and policies affect their customers (Gore, 1993, p. 108).

There has, however, long been another group of analysts who have insisted on the importance of a service output concept which is, like the concept of a physical good output, conceptually distinct from the outcomes (or utility) which result from the service. Today, this view is clearly in the ascendant. Typically, outputs are in contemporary usage defined as "the products and services produced by a program or activity" (World Bank, 1998, p. 133)[7]. In this usage, outputs "can never be outcomes", but are instead "the means by which outcomes are achieved" (Western Australian Department of Premier and Cabinet, 2003, p. 4).

This latter concept of a service output is clear and consistent with basic economic theory. It might be expected that the widespread adoption in a public sector environment of this concept of outputs would have eliminate the problem of distinguishing outputs from outcomes. This is not, however, entirely the case. There is a lingering ambiguity about the outputs/outcomes distinction, which manifests itself most notably in the

approaches which some analysts and practitioners take to the definition of completed *units* of output.

### *Proximate Outcomes, Quality and the Definition of a Unit of Output*

At what stage can the delivery of the service be considered to be complete? In other words, what constitutes a finished product?

One common approach is to define a completed unit of output in terms of the achievement of some result. Educational outputs are, for example, defined by some as "students graduated or promoted *who have met a minimum pre-specified standard of achievement*"[8]. Similarly, health outputs have been defined as "*effective* medical treatments" (HM Treasury, 2001, p.8, italics added). Such a definition of a unit of output means that the service-provider is only considered to have produced an output if a defined *proximate* outcome[9] of the service is achieved. For this reason, this type of definition of a unit of output will be referred to in this paper as the *successful output* definition.

If one adopts the successful output definition, then the distinction between an output and a *proximate* outcome becomes blurred – they are more or less two sides of the same coin. It is then only a short step further to *equating* outputs with proximate outcomes, as the economist T.P. Hill (1977, p. 318) did in an influential paper when he

defined an output as "the change in the condition of [the immediate] person or good affected" by the service[10].

This type of conflation of the concepts of outputs and proximate outcomes perhaps helps to explain the continuing contemporary incidence of output definitions which sound more like definitions of proximate outcomes. To give one recent example, the UK National Audit Office in its excellent report on *Measuring the Performance of Government Departments* makes the mistake of defining outputs as "the immediate result of Government activities", by contrast to outcomes which are defined as "the ultimate impacts on, or consequences for, the community of the activities of the Government" (2001, p 6).

Another approach which is sometimes used to defining a unit of output is to say that a service only becomes an output if and when it meets some specified quality standard or criteria. However, this approach often ends up being indistinguishable from the successful output definition. Thus, for example, when the US GASB (2003) formulated the definition of educational outputs quoted above, it suggested that the requirement of a "minimum pre-specified standard of achievement" was a "quality requirement". Once again, however, this blurs the concepts of outputs and outcomes – a fact of which the GASB seems to have been aware when it noted that "in some cases, meeting a quality requirement may turn an 'output' indicator into an 'outcome' indicator". This, however, raises a further significant issue: the relationship between the

concepts of *quality* and *outcomes*. Are they simply synonymous terms, as the GASB appears to imply?

The appeal of the successful output definition – and of the related quality standard output definition – is easy to understand. There is an understandable concern that if output quantity measures don't recognize quality or success, agencies might be able to 'get away' with delivering shoddy services while appearing to be performing well. Such enhanced output measures are particularly appealing if one is seeking to use output cost information to link relate budgets to output quantity targets, or to 'pay' agencies a 'price' for their outputs under so-called *output-purchase* budgeting systems (Robinson, 2002a, 2003). Notwithstanding this appeal, it is argued in this paper that these output definitions are generally incorrect, and that their use can have undesirable behavioral implications. It is a mistake to believe that output quantity measures can do double duty, as measures of outputs and outcomes. Outcomes must be measured directly.

It is necessary, in developing these propositions, to reconsider the relationship of outputs to outcomes. However, before doing so, it is useful to start by briefly clarifying two other areas where confusion often arises as to the meaning of the output concept. These concern, respectively, the relationship between outputs and support services, and the relationship between outputs and activities.

***Outputs and Support Services***

Consider an automobile company. The output – that is, the *product* – of the company is automobiles. No-one would consider the services provided by, say, the accounting department of the factory, or the human resources section, to be the company's products. Rather, they are support services – or, expressed differently, intermediate services in the process of car production. Yet it is not uncommon in the public sector for support services to be characterized as outputs. To pick merely one example, a recent report the US Office of Personnel Management (2001, p. 13) described as outputs the following: "files that are orderly and complete" and "a report that is complete and accurate".

If we wish to focus upon results rather than internal services and activities, it is preferable to reserve the term *output* to refer exclusively to services (or goods) delivered by an agency *to an external client or subject*. As the Australian Department of Finance puts it, "outputs are the goods and services produced by agencies on behalf of the Government for external organizations or individuals" (DOFA, 2003, p. 6). Services delivered to intra-agency clients should, by contrast, be referred to as support services[11].

### *Outputs versus Activities*

The definition of educational outputs as "students graduated or promoted who have met a minimum pre-specified standard of achievement" may raise the issue of the

appropriateness of the *successful output* definition of a unit of output, but there is one element of this definition which is indisputably correct. This is that it defines the educational output in terms of the service delivered to a student, rather than in terms of the *activity* of teaching. One nevertheless not infrequently comes across examples in official performance reports of supposed output measures which are in fact measures of service delivery activity. A good example arises in relation to the service that police provide when carrying out patrolling activities, the principle intended outcome of which is to deter crime. It is common for hours of police patrolling to be used as a supposed measure of output of such so-called *proactive* policing. However, patrol hours is clearly in fact an activity measure[12], and to use it as an output measure is akin to treating hours of bread-baking, rather than the bread itself, as the output of a bakery.

The erroneous identification of activities and outputs is, on the face of it, somewhat surprising, because at a conceptual level it might be thought that the distinction is reasonably clear. This confusion certainly cannot be attributed to the current enthusiasm for activity-based costing (ABC). The essence of activity-based costing is the use of activity cost pools and activity cost-drivers in order to cost outputs. ABC methodology is thus founded on a clear distinction between activities and outputs[13].

The main source of this confusion is the measurement difficulties which affect some outputs. It is difficult or impossible in the case of certain types of public sector services to observe and measure the numbers of subjects/clients to whom the service is delivered. To take the case of pro-active policing, patrolling achieves its deterrence

outcome tby creating an awareness of the police presence on the part of people in the patrol vicinity. Output quantity might, therefore, be appropriately b conceptualized as the number of people who become aware of the police presence. The obvious problem, however, is that measuring this output is clearly impractical. It is from this practical consideration that the need to use activity measures such as patrol hours arises[14]. There are other similar public services the outputs of which cannot be measured, and for which as a result of which it becomes necessary to place particular reliance upon activity measures. However, it is important that this pragmatic consideration should not lead us to confuse the output and activity concepts.

Does the distinction between outputs and activities matter? Given that services do not produce a tangible physical output, it could be argued that the activity and the receipt of the activity might appear to be merely two sides of the same coin, so that the distinction an immaterial one. This is, for example, true of the distinction between cutting hair (the *activity*) and receipt of a haircut (the *output*).

There are, however, a number of reasons why the distinction between outputs and activities does matter. Firstly, the production of an output commonly requires not one but a *set* of different, coordinated activities. Thus, for example, operative treatment in a hospital requires surgery, nursing and anesthesia.

Secondly, in the case of *collective* services, the fact that one set of activities yields multiple outputs makes the distinction vital. Education is, as indicated above, a good example – the activity of one teacher delivers an output to a number of students[15].

Thirdly, if one accepts the distinction between outputs and support services, then it follows that some activities produce intermediate service rather than outputs. ABC methodology distinguishes, for precisely this reason, between *primary* and *secondary* activities. Primary activities are activities which involve direct interaction with the (external) client/subject, whereas secondary activities do not.

## The Successful Output Concept

Against this background, we now return to the question of the appropriateness of the *successful output* definition of a unit of output.

As intimated above, the inclusion of an outcome requirement in the definition of a unit of service output implies a definition of service outputs which is very different from that used in conventional economic analysis for physical goods. Physical goods outputs are not defined in terms of outcomes (or utility) achieved. Unquestionably, a physical good has to be *capable* of producing the required outcome when it is used. However, this does not mean that an individual product only acquires its status as an output if and when it is successfully used to produce an outcome. No-one would suggest that a book only

becomes a book if and when its purchaser gets around to reading it. Similarly, a nail does not, for example, cease retrospectively to be a nail because the carpenter hits it at the wrong angle and bends it so badly that it can only be thrown away.

There is a superficially persuasive line of argument which suggests that the story is different for services. Given that services act directly upon a client or subject with the aim of effecting a proximate outcome, surely it is reasonable to say that the service output is only completed when the intended proximate outcome is achieved? A haircut, for example, can only be said to have been delivered when the customer's hair has actually been cut. Likewise, a painter cannot be said to have finished the job until the house he was hired to paint has actually been painted.

It is precisely this type of reasoning which underpins Hill's definition, referred to above, of outputs as proximate outcomes. He reasoned that services must involve:

> *some change … in the condition of some person or good…. Whatever the producer of the service does must impinge directly on the consumer in such a way as to change the condition of the latter. Otherwise, no service is actually provided. The mere performance of some activity is not enough if the consumer unit is not affected in some way.* (Hill, 1977, p. 318)

This led Hill to assert, by way of example, that the output of educational services is "the additional skill or knowledge imparted in a pupil directly as a result of the instruction

provided by a teacher" – so that if the pupil learns nothing, no output has been delivered (1977, p. 323-24).

It is, however, hard to be comfortable with this view. Consider two school students who sit in a class together during the year and receive exactly the same instruction. One does brilliantly while the other scarcely listens to the teacher and as a result either learns little or nothing. Is it appropriate to say that only in the case of the high-performing student has the teacher delivered an output? Note that if one's answer to this question is 'yes', one must be willing to attach the 'low productivity' label to teachers whose classes are filled with disadvantaged children who achieve lesser educational outcomes than more advantages children.

Consider, as a further example, the medical services provided to seriously injured patients admitted to the emergency unit of a hospital. The minimum desired proximate outcome will be to save their lives. However, even with the best care available, some of these patients will die. Is it really appropriate to assert that only those patients who are saved have received an output?

The underlying problem here is a very familiar one: the uncertainty and the variability which often characterizes the relationship between activity and outcomes. This, of course, reflects the mediating effect of *contextual factors*[16] – characteristics of the client/subject or of the service-delivery environment which mean that the same output may yield very different outcomes. This uncertainty and variability tends to be greatest

in relation to higher-level outcomes, but in the case of quite a few outputs it also affects the realization of proximate outcomes. Many of the services provided by government (or, indeed, purchased by consumers in the marketplace[17]) are desired by potential beneficiaries notwithstanding the fact that there is no certainty as to the extent to which, if at all, the desired proximate outcome will be achieved in any specific instance. Those beneficiaries will not want these services unless they believe that they are *capable* of yielding the desired outcome. But this does not necessarily mean that the outcome will in the event be achieved.

It follows that, while it is appropriate to hold that a service can only qualify as an output if it is potentially capable of producing a desired outcome, it is inappropriate to insist that outputs are produced only in those cases where an outcome is actually realized.

It is indeed partly for this precise reason that, as Burkhead and Miner (1971, p. 301) put it, "economic analysis has stubbornly eschewed measurement of quantities of goods in terms of their consequences". The economic concept of an output is entirely independent of the contextual factors which affect the utility/outcome which the output generates for the individual consumer. If this were not the case, and some utility or outcomes test were imposed upon physical goods, we would have an analytically impossible situation where, for example, two physically identical nails might no longer be considered to be the 'same' product.

The successful output definition of the unit of output is, in this respect, even more problematic than Hill's approach of equating outputs with proximate outcomes. Both are really outcome rather than output measures. But the successful output definition is inferior to Hill's approach even as an outcome measure. Hill's approach at least counts *any* proximate outcome which is delivered by a service. By contrast, the successful output definition is an outcome measure which disregards, firstly, any client outcomes below the level of the outcome standard and, secondly, any surplus outcomes above the outcome standard level. Thus, for example, once one has set some educational achievement standard for the purposes of a successful output measure, any student educational outcomes below that standard are not counted. Nor will there be any recognition of student educational outcomes *in excess* of the educational achievement standard. This is particularly problematic given that any educational achievement standard used for these purposes will be essentially arbitrary.

This analysis suggest that, in general, as outcome measures, measures such as average outcome and outcome variance (not to mention more sophisticated measures) will be superior as outcome measures to the successful output measure.

### *Defined-Outcome Services*

Before moving on to the question of quality standards, it is important to note one qualification to the above analysis. There is one sub-set of services for which it is

reasonable to apply a successful output test in defining units of output. Services such as hairdressing and house-painting are representative of this sub-set. For such services, it is understood that service activity will continue until that proximate outcome has been achieved – that, in other words, the output will only be considered to have been delivered when that point has been reached. Services where it is understood that service activity will continue until a defined proximate outcome is achieved will be referred to in this paper as *defined-outcome services*. In the tax-financed core public sector, such service may generally be said to have a *policy-based* outcome standard – in the sense that the setting of a pre-defined outcome standard will generally need to reflect a policy decision taken at some level within government. Whether provided in the public or private sector, services which fall into this category are for the most part characterized by the existence of little uncertainty about the extent of service activity required to yield the desired proximate outcome.

The fact that the successful output definition is appropriate for defined-outcome services does not, however, mean that it constitutes a suitable *general* definition of what constitutes a unit of output. For very many of the services delivered by government, it is *not* the case that the level of service activity applied to a specific client/subject in determined by what is required to achieve a pre-specified proximate outcome. For such services, any outcome standard used in a successful output definition must be inherently arbitrary.

## Quality and Activity Standards for Units of Output

As mentioned earlier, another approach to the definition of a completed unit of output is to apply a quality standard or criteria. If the supposed quality standard is outcome-based, this approach becomes simply the successful output definition in another guise. If, on the other hand, it is considered inappropriate to equate outcomes and quality, and some other concept of quality is employed, then the use of quality standards might represent an alternative approach to the problem of defining completed units of output.

On the first point, it is surely fundamentally mistaken to equate the quality of a unit of output with the particular outcome (or utility) which that output yields the individual consumer. No-one would, for example, impugn the quality of a meal produced by a top chef simply because it is not to the taste of a particular diner – particularly if, say, it turns out that the diner's culinary tastes were molded at McDonalds. What, then, is the relationship between outputs and outcomes? Outcomes are changes brought upon people, social structures or the physical environment. Output quality, by contrast, may be defined as to the extent to which the characteristics of the product are such as to increase its potential capacity to generate the desired outcomes[18]. For goods, characteristics means physical characteristics. For service outputs, characteristics can be viewed as referring to the quantity, standard, mix and delivery of the activities which are delivered to the client/subject in the course of the production of the service output (see Robinson, 2002a). Crucially, this is a concept of quality which captures not only *how* a service is

delivered, but *what* is delivered[19]. Defined in this manner, the quality of an output does not guarantee the achievement of an outcome in any specific instance. It simply makes the output more *capable* of producing the outcome. From this perspective, it is a mistake to confuse outcome standards and quality standards.

Having distinguished outcomes and quality, we are left with the question of whether it is appropriate to apply a quality standard proper to the task in defining completed units of output. The question becomes an intensely practical one: how are we going to operationalize the quality concept? The only obvious possibility involves the application of what is known as the *conformance to specifications* concept of quality. This involves pre-defining a set of characteristics which the output must possess if it is to be considered to be a completed product. This is a concept of quality particularly suited to mass-produced physical goods, where quality-control means judging each product against a standard checklist before giving it the 'OK' for sale. Given that services do not have a physical existence, applying this quality concept to services mean applying an *activity* standard. Concretely, the conformance to specifications approach implies that a service would be counted as a complete unit of output *only* if a certain pre-defined set of activities, appropriate for the type of service, had been delivered to the client/subject of that service[20]. We will refer to this approach as involving the use of an *activity standard* to define completed units of output.

The use of an activity standard in defining a unit of output is entirely appropriate for some services. It is particularly relevant to *standardized services*. These are services

where, as a matter of policy, the aim is that each client/subject receives the *same* pre-defined set of primary activities (standardized services could therefore also be referred to as *defined-activity services*). An example of a highly standardized service is motor license testing. Each motor license candidate will be tested in the same, or a highly similar, manner – through, perhaps, the administration of a standard written test followed by a practical driving test in a prescribed format. In the case of a standardized service, the pre-defined list of primary activities defines both a minimum *and maximum* level of activity which each client/subject will receive. Standardized services are thus also *uniformly-provided* services. (Note, in passing, that for some types of service it will make more sense to specify only a pre-defined *minimum* set of activities which will be delivered in every instance, leaving open the possibility that additional activities should be delivered to some client/subjects.)

There is a crucial difference between *policy-based* activity standards and *arbitrary* activity standards. In the former case, there is an explicit policy (or service standard internalized by the professional training of the relevant staff) requiring that all clients/subjects receive *at least* the set of primary activities which comprises the activity standard. If there is no such policy, then any activity standard used for output definition purposes would be inherently arbitrary.

***Limits on the Scope for the Use of Activity Standards***

Useful as activity standards are for some services, there are limits upon the extent to which it is possible to define policy-based activity standards for tax-financed services.

One reason for this is that it is in the case of some services desirable to to vary the activity content of outputs significantly over time in order to deal with variability in need or demand. In the case of outputs which are sold on the market, if demand for the product increases, production will be ramped up because extra sales will cover the extra production costs. In the public sector, by contrast, agencies face relatively inflexible hard budget constraints, arising ultimately from the fact that their output is tax-financed rather than sold. Under these circumstances, increased need/demand will often – particularly if unanticipated – have to be dealt with by tighter rationing of services. Rationing could take the form of quantitative rationing, meaning that a greater number of clients/subjects are screened out, with a consequent increase in unsatisfied demand or lengthened waiting lists[21]. Alternatively, it could take the form of qualitative rationing, involving a reduction in the activity delivered to the average client/subject in order to permit the more clients/subjects to be handled with the same resources. Qualitative rationing is used extensively in the public sector, not only because it is in some cases an appropriate form of rationing from a welfare perspective, but arguably also because political considerations make politicians overly averse to quantitative rationing. In the case of a service subject to significant demand/need fluctuations, there is no point basing the definition of a completed unit of output upon an activity standard unless the policy is that each client/subject will receive at least the minimum set of activities embodied in the activity

standard *even* if there is a substantial unanticipated demand surge requiring the use of quantitative rationing.

A further reason that output-unit definitions based upon activity standards may be inappropriate is the prevalence within the public sector of output *heterogeneity*. Heterogeneity refers to deliberate variations in the set of activities delivered to clients/subjects in response to differences in client/subject characteristics. In the case of heterogeneous services, there is no intention that the service be uniformly-provided. Examples of heterogeneity, and a discussion both of the relationship of heterogeneity to output classification and of some of the implications of heterogeneity, may be found elsewhere (Robinson, 2002a).

The fact that there may be significant variation in the characteristics of client/subject of a particular type of output does not necessarily mean that the output will be a heterogeneous one. Heterogeneity requires a decision to tailor the output in response to those varying client/subject characteristics. The decision on whether, and how far, to tailor the output will reflect, in part, what Sen (1992) calls the operative concept of "basal equality". For some types of service, equality in the level of service activity provided to clients may be particularly valued, in which case the service will tend to be uniformly-provided. For other services, however, greater weighting may be placed on equality of outcomes. In this latter case, if there is considerable variation in relevant client/subject characteristics, the service will be characterized by significant output heterogeneity, the aim of which is to achieve a less unequal distribution of outcomes than would arise from

uniform provision. Generally, however, the tailoring of service provision in order to address concerns about outcome inequalities will not, or cannot, go so far as to bring about full equality of outcomes[22]. Thus heterogeneous services will generally not fall into the category, discussed above, of defined-outcome services.

The desire of policy-makers to reduce inequality of outcomes is not the only source of output heterogeneity. Heterogeneity may also arise from a desire to maximize the overall effectiveness of a program. A simple analogy with welfare economics tells us that maximizing the aggregate outcome delivered by a service, subject to a given budget constraint, would require that the activity set delivered to each client/subject be varied so as to equalize marginal (expected) outcomes per dollar for each individual. Such effectiveness-driven heterogeneity might well *increase* inequality of outcomes relative to uniform provision[23].

## Units of Output: a Suggested Definition

The above analysis suggests that in the absence of a pre-defined policy requiring that each client/subject receive at least a certain minimum set of service activities, it is inappropriate to define completed units of output by reference to an arbitrarily-specified activities standard. It has also been argued that, unless there is a pre-defined policy requiring that activity continue to be delivered to the client/subject until a pre-defined proximate outcome is achieved, it makes little sense to use a successful output definition

of the unit of output based upon an outcome standard which will necessarily be arbitrarily in nature.

There are probably many public sector services for which it is not feasible, or appropriate, to set either outcome standards or activity standards. Unfortunately, one implication of this is that, for some services, the concept of a completed unit of output – and the distinction between an incomplete and a completed product – will unavoidably lack precision. For such services, the most that it may be possible to say is that, in order to qualify as a completed unit of output, the set of activities delivered to the client/subject must have been sufficient to have been *potentially capable* of producing some positive level of a desired proximate outcome. To this extent only can we agree with Hill (1977, p. 318) that the "mere performance of an activity" cannot constitute an output.

This analysis points towards the following definition of a completed unit of output:

*An output is produced when there is delivered to a client/case a set of activities which is considered to be potentially capable of to some degree inducing a desired outcome, subject to two further provisos: Firstly, the activities delivered must include any minimum set of activities which policy-makers have determined in advance will, irrespective of workload pressures or differences in client/subject characteristics, be delivered to all clients/subjects – if such a set of mandatory activities exists. Secondly, if and only if policy-makers have determined that activities will continue to be delivered to*

*the client/subject at least until the point where a pre-specified proximate outcome is achieved, an output will not be considered to have been produced unless that proximate outcome is in fact achieved.*

## Behavioral Implications of Inappropriate Definition of the Output Unit

As noted earlier, it is increasingly commonplace these days for funding, or other rewards/penalties, to be linked to agency performance indicators. It is generally well understood that, if funding or other rewards/penalties are linked to outputs, the definition of outputs which is employed may have significant behavioral implications. It is useful therefore to examine briefly some of the possible behavioral implications of inappropriately employing either the successful outcome or activity standard definition of the output unit. It is convenient to do this by considering circumstances where agency funding is based directly upon the payment of a 'price' per output[24].

As noted above, the use of an arbitrary successful outcome definition of outputs (i) disregards any client outcomes in excess of the outcome standard employed and (ii) gives no recognition to levels of outcome achieved which are below the outcome standard. It follows that if this output concept were to be used as the basis of agency funding, it would encourage equalization of client/subject outcomes (at the level of the outcome standard). This tendency to the equalization of outcomes would arise from the fact that agencies would have an incentive to commit only the minimum amount of

activity to each client/subject necessary to achieve the outcome standard[25]. There is one significant qualification to this tendency for outcome equalization: namely, that paying only for those outputs which achieve the outcome standard may be expected encourage the denial or minimization of service to clients/subjects in respect of whom the expected total cost of achieving the outcome standard exceeds the price paid per output.

The key question which this raises is whether such a set of behavioral incentives is compatible with policy-makers' notions of "basal equality" – and with their view of the how far trade-off equality of outcomes should be pursued at the expense of the maximization of aggregate outcomes – for the service concerned.

Take the example of school education. There is ongoing active debate amongst educational policy-makers concerning the degree of tailoring (heterogeneity) which should characterize education. However, there is considerable broad agreement that their should be explicit service heterogeneity in relation to at least two groups of students, namely:

- students with serious educational disabilities, in relation to whom policy-makers would like to see additional service activity delivered in order to reduce somewhat the degree of outcome inequality these students would experience if subject to uniform provision.
- gifted students, in relation to whom policy-makers would like to see additional activity delivered because expenditure on such students is particularly cost-effective

in raising the aggregate effectiveness of the education system, and thereby its contribution to higher-level outcomes such as economic prosperity.

Suppose that schools were to be funded on the basis of outputs as defined along GASB lines ("students graduated or promoted who have met a minimum pre-specified standard of achievement"). The behavioral distortions induced by such an approach have been explored by Mayston (2002) in an analysis of the implications of aspects of the school funding model in Britain. As he points out, such a funding model encourages schools, contrary to the wishes of policy-makers, to neglect any student who they felt they could not without great assistance reach the operative educational achievement standard – including, most obviously, students with serious educational disabilities. In the second place, schools would devote less rather than more educational activity to gifted students than to average students – because it requires less effort to teach a gifted student to achieve any specific level of educational achievement. Thus the incentives created by funding based upon the successful output concept would be incompatible with both the equity and effectiveness preferences of policy makers.

The root of the efficiency problem arising from the successful output concept is obvious from the standpoint of economic analysis. It is that whereas the use of the successful output concept encourages considerable equalization of the outcomes achieved by students, the maximization of aggregate outcomes requires (approximately speaking) the equalizing of marginal outcomes per dollar.

The effects of the use for funding purposes of a definition of output units based upon an *arbitrary* minimum activity standard are even simpler. Such an approach would, firstly, create incentives for reduced heterogeneity, and towards a more uniformly-provided service. The other effect would be to promote greater reliance upon quantitative rationing in response to surges in need/demand. Again, this would in the case of many services conflict with the preferences of both policy-makers and the community.

## Conclusion

It is, in general, inappropriate to define a completed unit of output by reference to either outcome or activity standards. The use of such standards in output definitions is appropriate only in respect to those services which have been referred to as *defined outcome* or as *defined activity* (standardized) – services which constitute only a sub-set of the services provided by government. The application of essentially arbitrary outcome or activity standards to public sector services which are not part of this sub-set is not only inappropriate, but may lead to serious behavioral distortions. Rather than seeking to incorporate arbitrary outcome standards in output unit definitions, outcomes should be measured directly through measures which count all outcomes achieved rather than simply counting the number of times an arbitrary outcome standard is attained.

## References

Australian National Audit Office (1996). *Better Practice Principles for Performance Information*. Retrieved May 1, 2003 from http://www.anao.gov.au/WebSite.nsf/Publications/4A256AE90015F69B4A25691B001E1E9F.

Berman, E. M. (1998), *Productivity in Public and Non-Profit Organizations: Strategies and Techniques*. Thousand Oaks, California: SAGE Publications.

Culyer, A. (1983). *The Political Economy of Social Policy*. London: Martin Robertson.

Department of Finance and Administration (2003). *Aims User Manual*. *Module 1: Business Processes and Concepts*. Retrieved May 1, 2003 from http://www.dofa.gov.au/budgetgroup/AIMS__Accrual_Information_Mana/aims_-_material_agencies_manua.html.

Government Accounting Standards Board (2003). *The Elements of Performance Measurement Reporting***.** Retrieved February 28, 2003 http://accounting.rutgers.edu/raw/seagov/pmg/perfmeasures/index.html.

Gore, A. (1993). *Creating a Government that Works Better & Costs Less: The Report of the National Performance Review*. New York: Plume.

Hill, T. P. (1977). On Goods and Services. *The Review of Income and Wealth*, *23*, 315-338.

HM Treasury (2001). *Choosing the Right Fabric: a Framework for Performance Information*, London: Author.

Mayston, D. (2002, August). Efficiency and Effectiveness in Public Services: Outstanding Problems of Performance Measurement. Paper presented to the Conference of the International Institute of Public Finance, Helsinki. Retrieved August 22, 2003 from http://data.vatt.fi/iipf2002/members/papers.htm.

National Audit Office (2001). *Measuring the Performance of Government Departments*, London: The Stationery Office.

Propper, C. (2001). 'Expenditure on Healthcare in the UK: a review of the issues'. *Fiscal Studies*, *22*, 151-183.

Robinson, M. (2002a). Output-Purchase Funding and Budgeting Systems in the Public Sector. *Public Budgeting and Finance*, *22*, 17-33.

Robinson, M. (2002b). *Best Practice in Performance Budgeting*. (Economics, Finance, and International Competitiveness Discussion Paper no. 124). Brisbane, Australia: Queensland University of Technology, School of Economics and Finance. Retrieved August 22, 2003 from http://www.bus.qut.edu.au/schools/economics/research/disc_pre2001.jsp.

Sen, A. (1992). *Inequality Reexamined*. Oxford: Clarendon Press.

Simpson, W. K. & Williams, M. J. (1996, Spring). Activity-based Costing, Management and Budgeting, *The Government Accountants Journal*. 10-14.

Stiefel, L., Rubenstein, R. & Schwartz, A. (1999). Using Adjusted Performance Measures for Evaluating Resource Use. *Public Budgeting and Finance*. *19*, 67-87.

United States Office of Personnel Management (2001). *A Handbook for Measuring Employee Performance: Aligning Employee Performance Goals with Organizational Goals*, PMD-13.

Western Australian Department of Premier and Cabinet (2003). *Preparing Performance Indicators - A Practical Guide*. Retrieved May 1, 2003 from http://www.dpc.wa.gov.au/index.html.

World Bank (1998). *Public Expenditure Management Handbook*. Washington, DC: Author.

## Endnotes

[1] Other synonymous or closely-related terms (such as 'end outcome', 'low-level outcome', 'intermediate outcome') are often used.

[2] What is referred to here as the *subject* of a service may be human (e.g. a student), or it may be physical (e.g. a building on fire). A human subject may or may not be an intended beneficiary of the service (consider, for example, a student versus a prisoner). For that reason, it is arguably inappropriate to refer to *all* of the human subjects of services as 'clients'. The term is more suited to human subjects who are also intended beneficiaries of the service concerned.

[3] Greater precision might be gained by referring to the intended changes brought about in the subjects of services as 'direct outcomes', where direct outcomes are one specific type of proximate outcome. The term is, however, not part of the conventional performance lexicon and will not for this reason be used here.

[4] Culyer sometimes used the term 'throughput' to mean, approximately, what is in contemporary parlance an output. Some others today use the term 'workload' in the same manner (e.g. Berman, 1998, p. 59)

[5] For example, Propper (2001, p. 171-2) refers to 'health' as the 'output' of the health system.

[6] Thus Berman (1998, pp. 53-7): "outputs are defined as outcomes that are direct, immediate consequences of strategies …For example, teacher outputs include student test scores, graduation rates …".

[7] Similar contemporary examples of the definition of outputs are: "the products or services which are produced and delivered by a program" (ANAO, 1996, p. 33) and "outputs measure the quantity of services provided" (GASB, 2003).

[8] This wording is attributable to the US GASB (2003). Italics added. See, similarly, Stiefel, Rubenstein and Schwartz (1999, p. 69), who assert that the educational output is the "number of graduates".

[9] Once again, we could more precisely talk about a defined *direct* outcome upon the subject being achieved.

[10] What we might call, if being completely precise, the *direct* outcome upon the subject of the service.

[11] This leaves unresolved the more moot point as to whether services provided by one government agency to another (rather than to a client/subject external to government) should be regarded as outputs. Space restrictions prevent the discussion of this point here.

[12] Similarly, in the Australian national government's performance concepts framework, the term 'activity' is used to refer to disaggregated classes of outputs.

[13] Not that this prevents confusion on the part of some, as exemplified by the following: "Activity-Based Costing provides a technique for cost control that assigns costs – both direct and indirect – of products or services (activities), based on the consumption of resources….. ABC computes the total cost of activities (products) as well as the unit cost of activity components, sub-activities, tasks and phases. ….Output measure = Quality of work accomplished." (Simpson and Williams, 1996)

[14] The police patrolling example raises another issue, which can only be touched on here. This is the indeterminacy of the output/cost relationship for 'collective' services – services were one (set of) activity yields multiple outputs. This arises from the fact that collective services' possess the public goods property that, at least within some range, the marginal cost of additional units of output is zero.

[15] It is also, in the case of collective services, possible to have the activity without the output – consider, for example, an after-hours bus service which nobody boards on a particular night.

[16] Sometimes also referred to by the less neutral term "confounding factors".

[17] If, for example, someone charged with a criminal offence hires a lawyer to defend himself, there is no certainty that the defense will succeed.

[18] This is, of course, what is known as the 'fitness for purpose' concept of quality.

[19] It is thus a somewhat broader concept of quality than that put forward by the World Bank in its excellent *Public Expenditure Management Handbook,* where it is suggested that quality 'refers to characteristics of *how* a product or service is delivered' (although it should be noted that the Bank clearly effectively recognizes that quality relates to *what* is delivered when it observes that 'conformance to specifications' is one possible quality characteristic). Arguably, such a definition of quality is much more suited to standardized services than to the heterogeneous services which are so commonplace in the public sector. Note also that although the World Bank correctly points out that quality does not refer to the 'results of the service', it then makes the conflicting statement that quality is an intermediate outcome (1998, p. 134). From the perspective advanced in this paper, the fact that output quality is a means of achieving outcomes does *not* make it an intermediate outcome.

[20] In this context, 'set' refers to both the mix of different activities and the quantity of each activity.
[21] A standardized service will by definition be subject to such a quantity rationing response if there is a surge in demand without a commensurate increase in resources.
[22] For some services, the cost of achieving equal outcomes for the more difficult clients/cases will be regarded as unacceptably high. For other services (e.g. many medical treatments) it will be impossible to guarantee equal outcomes no matter how much activity is lavished on the more difficult clients/cases.
[23] Other factors which may influence the degree of heterogeneity in the provision of a particular service include: the underlying degree of variation in relevant client characteristics/contextual factors, which determines the extent to which there is conflict between the goal of outcomes equity and the maximization of aggregate outcomes; the indivisibilities which characterize some outcomes (e.g. it is not possibly to partly save someone's life in an emergency ward, whereas schools can certainly 'partly' educate students); and concern with level and distribution of higher-level outcomes.
[24] Such an 'output-purchase' budgeting system is quite a long way removed from the complexities of real-world public budgeting, but such a simple model does have the advantage of putting to spotlight on the behavioral effects of output unit definitions.
[25] And only if they then have 'surplus' resources to consider the distribution between clients of additional activity