

# **The Meta-evaluation of the Sports Participation Impact and Legacy of the London 2012 Games : Methodological Implications**

[Paper accepted with both forms of abstract 9 March 2016 for publication in Journal of Global Sport Management]

***Ian Henry, Centre for Olympic Studies & Research, Loughborough University***

## **Short Abstract 1**

The London 2012 Games were subject to the most substantial evaluation of any of the Olympic Games, or indeed any other sporting event, to date, in the form of a meta-evaluation. Meta-evaluations evaluate the collective lessons learned from smaller individual evaluation studies and consist of two principal elements. The first reports a synthesis of the results of the evaluations of individual programs and projects, and of national participation data; and the second consists of an evaluation of the rigour of the methods adopted in project and program evaluation and national evaluation studies, and of the conclusions drawn

The paper addresses one of the high-profile legacy goals associated with the 2012 Games, increased sports participation, and seeks to explain the disjunction between national level participation data indicating little or no increase in participation and programme or project data suggesting significant increases, and thus highlights the limits to this specific application of the meta-evaluation approach.

## **Long Abstract 2**

The London 2012 Games were subject to the most substantial evaluation of any of the preceding Olympic Games, or indeed any other sporting event, to that date. This evaluation took the form of a meta-evaluation. Meta-evaluations evaluate the lessons learned from smaller individual evaluation studies and consist of two principal elements. The first reports a synthesis of the results of the evaluations of individual programs and projects, and of national participation data; and the second consists of an 'evaluation of evaluations' which undertakes a critical review of the methods adopted in those project and program evaluation and national evaluation studies, and evaluates the rigour of their implementation.

The paper addresses one of the high-profile legacy goals associated with the 2012 Games, that is the claim that staging of the Games would inspire greater participation in sport and physical activity on the part of the host country population. A range of public, voluntary and commercial sector bodies initiated undertakings aimed at fostering and leveraging increased sport participation, from individual projects to large scale programmes. Using data from two national surveys of sports participation (the Active People Survey, and the Taking Part Survey, respectively commissioned by Sport England and by the Department of Culture, Media and Sport), and data from all the available evaluation studies of individual

projects and programs initiated in connection with planning associated with the staging of the 2012 Games, the paper identifies conflicting conclusions drawn from these two sources. The former (national survey data) indicated that little or no increase in sports participation had taken place in the run-up to the 2012 Games (with the exception of one period early in 2012 itself); while the latter (program and project data) suggested significant increases in forms of participation.

To explain this phenomenon the paper highlights key problems with the quality of data at the national survey level (shortcomings in the methodology); problems at the program level (issues of validity and reliability in defining and calculating participation); with the quality of evaluations (and in particular the lack of calculation of additionality); with the lack of evaluation per se in many instances where only crude participation data for projects and programs were available); and with the lack of focus on realist evaluation of the causal factors associated with the outcomes to be achieved, which would allow explanation of the achievement or failure to achieve increased participation patterns.

While the meta-evaluation approach represented a significant step forward in terms of the sophistication of methods employed in legacy evaluation of major sporting events, the approach is limited to some degree by the difficulties of combining results of local and national analyses, as well as of aggregation of results of evaluation studies that have adopted methods of varying types and which allow of variable levels of complexity of analysis.

## 1. The London 2012 games context

When London presented its bid to become the host city for the 2012 Games to the IOC Congress in Singapore in July 2007 it did so in a relatively daring manner focusing its presentation on three principal themes: inspiring youth (not simply in the UK but around the world); using sport for promoting intercultural understanding in particular through sport for development projects, but also domestically by celebrating the cultural diversity of the London Olympic Boroughs; and finally by reference to the use of sport for urban regeneration in a very disadvantaged part of the capital. These themes addressed outwards to the principal external stakeholders (in particular the IOC) represented a shrewd selection which played to the concerns of the Olympic movement concerning the decline in media interest detected among the young (which had been a stimulus to inaugurating the Youth Olympic Games), and was communicated through what was perceived as being a very effective and emotive presentation (Lee, 2006). Domestically, stakeholder concerns were different, with emphasis placed on urban regeneration; the promotion of national image through sporting success; facility legacy; and promotion of social inclusion (Department of Media Culture and Sport, 2008, 2010; UK Government & Mayor of London, 2013). These two sets of factors together were combined in the legacy goals subsequently proclaimed by government.

It is worth remembering that the Games Bid was won when Britain was still governed by a New Labour administration, and that the legacy claims were initially set by the Labour government in terms of 5 promises. Interestingly these promises had little to say about socially progressive goals other than referring to the regeneration of East London. In particular the government was criticised for not including a promise in relation to Paralympic sport and the position of those with disabilities (though the government responded by adding a legacy claim relating to Paralympic sport later in the day) (UK Government, 2015). Labour's five legacy promises were:

*Promise 1– Making the UK a world-leading sporting nation (legacies in (a) sport participation and (b) elite sport (c) Sporting infrastructure)*

*Promise 2– Transforming the heart of East London (regeneration legacy)*

*Promise 3– Inspiring a new generation of young people (both locally, and globally International Inspiration Legacy)*

*Promise 4– Making the Olympic Park a blueprint for sustainable living*

*Promise 5– Demonstrating the UK is a creative, inclusive and welcoming place to live in, visit and for business*

*(Adapted from; Department of Culture Media and Sport, 2008)*

Thus when the new Coalition government came to power in 2010 it probably had little leeway in terms of its ability to reshape the legacy aims of the Games and the Coalition's four Legacy themes were fairly close to those of its Labour predecessor.

- Sport: harnessing the UK's passion for sport to increase grass roots participation and competitive sport and to encourage physical activity;
- Economy: exploiting the opportunities for economic growth offered by hosting the Games;
- Community Engagement: promoting community engagement and achieving participation across all groups in society through the Games;
- East London regeneration: ensuring that the Olympic Park can be developed after the Games as one of the principal drivers of regeneration in East London. (Department of Culture Media and Sport, 2010)

However there are perhaps some subtle differences between the two statements of legacy aspiration. The first is the mention of community engagement by the Coalition (Ray, 2012) which could be interpreted fairly widely and might for example include engagement of those with disabilities. It might also relate to the theme of 'Big Society' promoted by the Conservatives, the dominant party in the coalition, which sought to engage individuals and groups in providing contributions to society that replace state provision (Featherstone, Ince, Mackinnon, Strauss, & Cumbers, 2012). The second is the emphasis placed on competitive sport for young people. In the cuts introduced in the first year of the Coalition government approximately £280 million was withdrawn from the funding of School Sport Partnerships with a sum of £10 million added to fund a UK Schools Games. The SSPs had been highly successful in generating additional PE experiences for children (SecEd, 2013).

## **2. Meta-evaluation as an approach to assess the Games legacy**

This paper will focus less on the overall assessment of the intended legacy gains and actual outcomes than on aspects of the methodology employed in the evaluation of the 2012 Games. The Labour government in effect made a bold decision to evaluate the achievement of its Olympic objectives by means of a comprehensive three-year Meta-Evaluation of the Impacts and Legacy of the 2012 Olympic Games and Paralympic Games commissioned by the Department of Culture, Media and Sport from a consortium of organisations led by the consultants Grant Thornton (the other partners being being Loughborough University, through its Centre for Olympic Studies and Research, the ECORYS consultancy, and Oxford Economics). The study in effect was not simply an investigation of the legacy outcomes for the 2012 Games, it was also used to evaluate the meta-evaluation methodology itself, with part of the study funded by the Economic and Social Research Council. As the authors of the methodological review of the meta-evaluation pointed out, the study was intended to be ground-breaking.

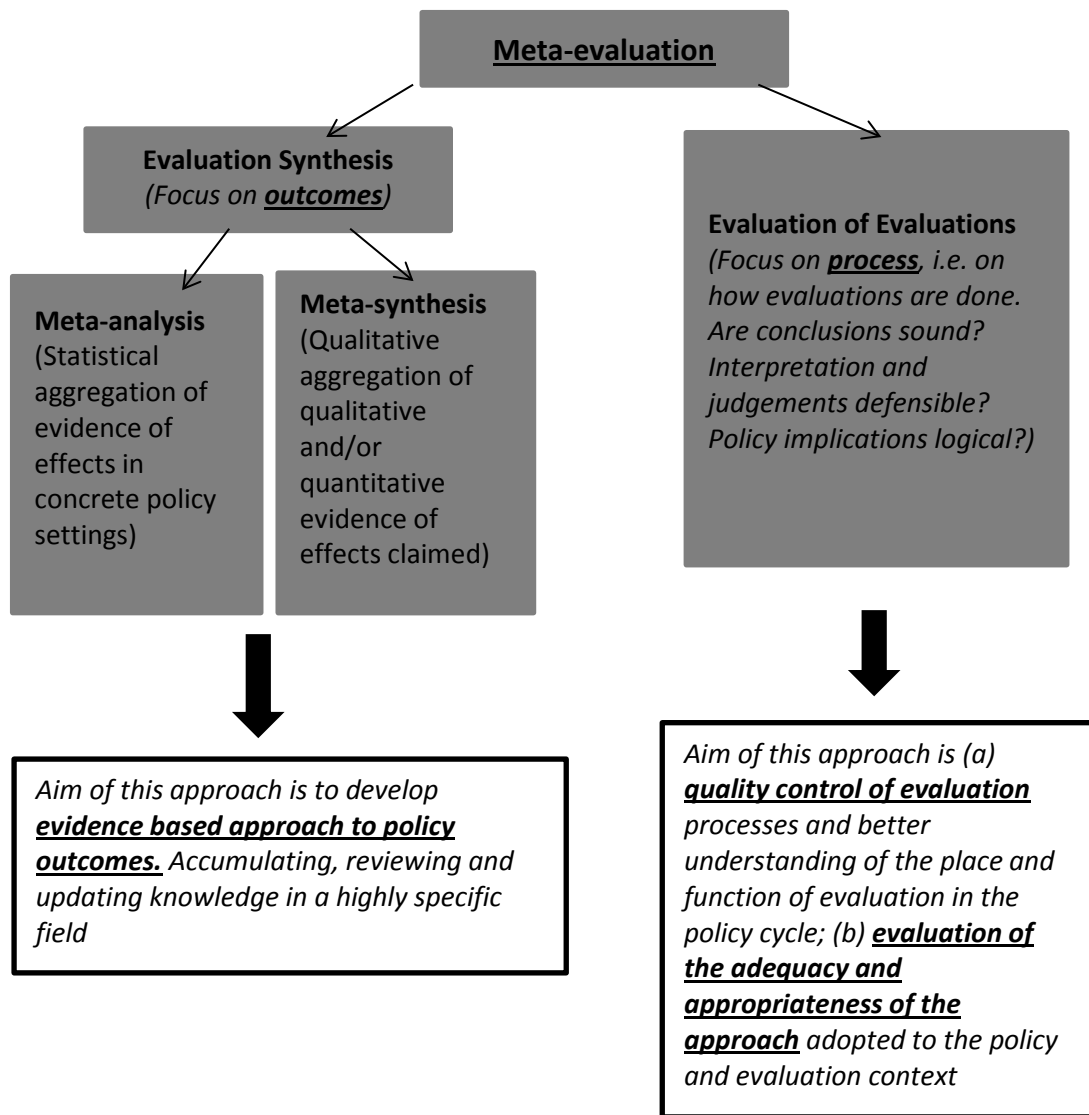
*The study is of the utmost importance in demonstrating the legacy impact of the 2012 Games across all thematic areas and will be the single largest and most comprehensive evaluation exercise commissioned in connection with the event (Gough, Martin, Grant Thornton, & ECORYS, 2012).*

We have sought to define the meta-evaluation approach elsewhere (Chen, Henry, & Ko, 2013) but it is worth briefly recounting the nature of the approach, its strengths and limitations. The term meta-evaluation may be traced back to the late 1960s (Scriven, 1969) though its application has evolved along two principal dimensions as illustrated in Figure 1, namely on forms of **evaluation synthesis**, combining the power of individual evaluations to draw more robust, evidence-supported conclusions; and the **evaluation of evaluations**, assessing the quality of methods adopted and thus the confidence that might therefore be placed in the findings of these studies.

Evaluation synthesis takes one of two primary forms, namely meta-analysis or meta-synthesis. Meta-analysis focuses on the combining of findings of independent studies which have employed statistical analysis of the effectiveness of particular interventions. It is most commonly applied in fields such as medical / health or psychological studies. Here the identification of appropriate studies addressing the impact of an intervention, is undertaken primarily through a Systemic Review of the literature (Petticrew & Roberts, 2006), identifying studies or trials which meet appropriate criteria of methodological rigour (typically in the medical sciences through randomly controlled trials). This is followed by analysis of the range of effect sizes of the studies included in the analysis, weighted according to the size of sample, to establish with a greater degree of precision what claims

can be made with reference to this wider meta-sample of studies of interventions. Meta-analysis of studies thus provides a more precise statistical estimate of 'treatment effect', than single studies will allow (Crombie & Davies, 2009).

**Figure 1: The Concept of Meta-evaluation**



Meta-synthesis of studies takes the same starting point of a Systematic Review, identifying qualitative and quantitative evaluations of the impact of a particular phenomenon or intervention. Identification of appropriate studies to incorporate in the synthesis depends on the establishing of criteria of quality of the studies under review. This is a much more complex and contested process than the meta-analysis approach which depends to a large degree on the RCT as the 'gold standard' methodology, and the application of standardised statistical protocols. As a consequence there is a much wider range of methodologies for meta-synthesis than is the case for meta-analysis. These include for example meta-ethnography (Noblit & Hare, 1988), thematic synthesis (Thomas & Harden, 2008), meta-narrative approaches (Greenhalgh, Wong, Westhorp, & Pawson, 2011), and textual narrative synthesis (Lucas, Arai, Baird, Law, & Roberts, 2007), which rely on a range of ontological and

epistemological assumptions (Barnett-Page & Thomas, 2009). What these approaches have in common is the focus on synthesising from qualitative analyses the conclusions which might be drawn through the pooling of evidence across studies of different types but with common standards of rigour applied.

The evaluation of evaluations is effectively a means of establishing the rigour with which evaluations have been undertaken and thus a means of establishing whether particular studies should be included in any meta-analysis or meta-synthesis.

### **3. Evaluating the claims that hosting the Games will result in higher levels of participation.**

In order to illustrate some of the difficulties encountered in applying the meta-evaluation approach we will focus on one of the problematic legacy goals that of using the games to leverage higher levels of sport and physical activity participation. Analysis of whether this goal was being achieved in the run up to the Games and the immediate aftermath was conducted on two levels. At the national level, England (rather than the UK) had two major national studies of sport and active recreation participation, namely the *Active People Survey (APS)* and the *Taking Part Survey (TPS)*, together with a national survey of sport and physical education associated with the *Physical Education and Sport Strategy for Young People (PESSYP)*, which provided cross-sectional data on the level of sports participation among the general population (for the APS and the TPS) and the school population (PESSYP). These 'top-down' national surveys were complemented at the project or programme level, by data from projects/programmes designed to stimulate increased participation, relating to uptake in those projects. This project / programme level evaluation provided 'bottom up' data on numbers engaging in these projects, complementing the 'top down' analysis derived from the national participation data.

#### **3.1 National Surveys of Sport Participation – Difficulties of Measurement**

The Active People Survey (Sport England, 2015a) and the Taking Part Survey (DCMS, 2012) involved data collection across the period from 2005 to 2012, and the data from both surveys provided a broadly similar picture of participation using different measures and methodologies. (From the end of 2012 sports participation data were to be reported only by Sport England in the Active People Survey, and were no longer analysed or reported on by DCMS or in the Taking Part publications.)

The Active People Survey (APS) had until 2012 used household landline telephone surveys on a sample of 161,000 households, while the Taking Part Survey (TPS) employed face-to-face interviews in the homes of interviewees. In 2012/13, 9838 adults and 801 children aged 11-15 were interviewed for the Taking Part Survey<sup>1</sup>.

In its written statement to the House of Lords Select Committee on the legacy of the 2012 Games, Sport England submitted the following in evidence

*Analysis of overall participation levels since 2005 shows a steady upward trend. The initial figure of 13.9 million people (34.2% of the population) for the period October 2005/6 had increased to record levels by October 2011/12, when it reached 15.5 million (36.0%)—1.6 million more people playing sport. The period between October 2010/11 and October*

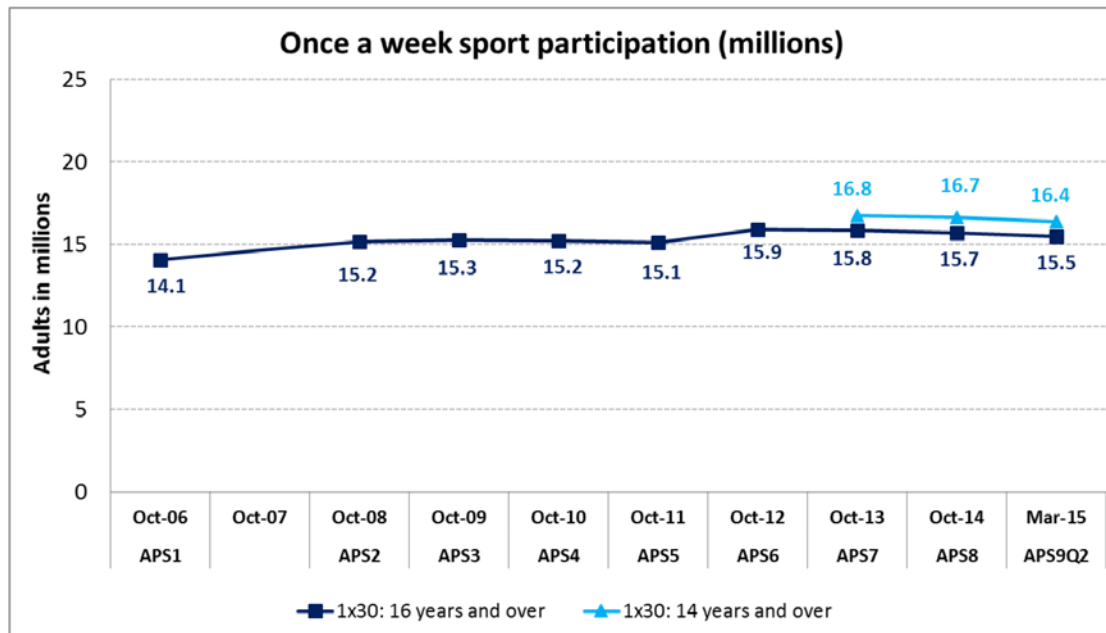
---

<sup>1</sup> Information was also collected in 2012/3 for the first time from parents or guardians of children aged 5-10 (n=1136) in relation to children's participation in sport.

2011/12 saw a significant increase of 753,600 people, with the majority of that growth (578,500) driven by women. The most recent figures, released in June 2013, showed that most, but not all, of that growth has been sustained. The current level of 15.3 million means that 533,000 of the 753,600 gained have been retained. While it was disappointing to see the slight dip in figures, it was not unexpected due the exceptionally cold weather in January and March. There is confidence among many sports that figures are already showing signs of recovery, suggesting that the dip in figures is temporary, and the longer term upward trend will continue. (House of Lords, 2013: ch. 2, para 115).

This conclusion appears over-optimistic in the light of subsequent data from the APS up to March 2015 illustrated in Figure 1, which shows a tailing off of participation from October 2012 to March 2015.

**Figure 1: Data Relating to Sport Participation from the Active People Survey (2006-2015)**



Source: (Sport England, 2015b)

The data from the APS however has been subject to criticism on a number of counts. It was recognized that the use of landline only phone connections skewed the collection of data towards the older end of the population<sup>2</sup> and the use of face-to-face interviews (as employed in the TPS) was deemed to provide more accurate data. The need for data relating to the younger population was also a limiting issue (data collected was for the adult population only, aged 16+, in APS until 2013).

In addition to the data for the period post 2012 pointing to difficulty in sustaining the statistically significant increase recorded in the peaking of participation in 2012, the gaps in participation between different age and socio-economic groups persisted, with younger groups and higher socio-economic groups experiencing greater growth from the first wave of APS data (APS1 in 2006) to the most recent (Quarter 2 of the ninth wave of data, APS9 in

<sup>2</sup> Ofcom research suggests that just over a quarter of 16-24 (26%) and 25 to 34 year olds (28%) lived in a household that used mobiles as its sole form of telephony in Q1 2014, significantly higher than the 16% average recorded across all adults during the period.

2015). Participation by black and ethnic minority groups however had grown more substantially than was the case for the 'white British' population across this period (Sport England, 2015b).

In the period running up to the Games a significant anomaly emerged in that at this stage (i.e. prior to 2012 itself) the national survey data was suggesting no statistically significant growth in participation, while data from national and local programmes were reporting major 'increases' in participation within these programmes. This disjunction between programme / project 'bottom up' data, and national participation survey 'top down' data may be explained in a number of ways, as we outline below.

### **3.2 Programmes and Projects aimed at Increasing Participation: Difficulties of Measurement**

#### **3.2.1. The quality of the data at programme / project level**

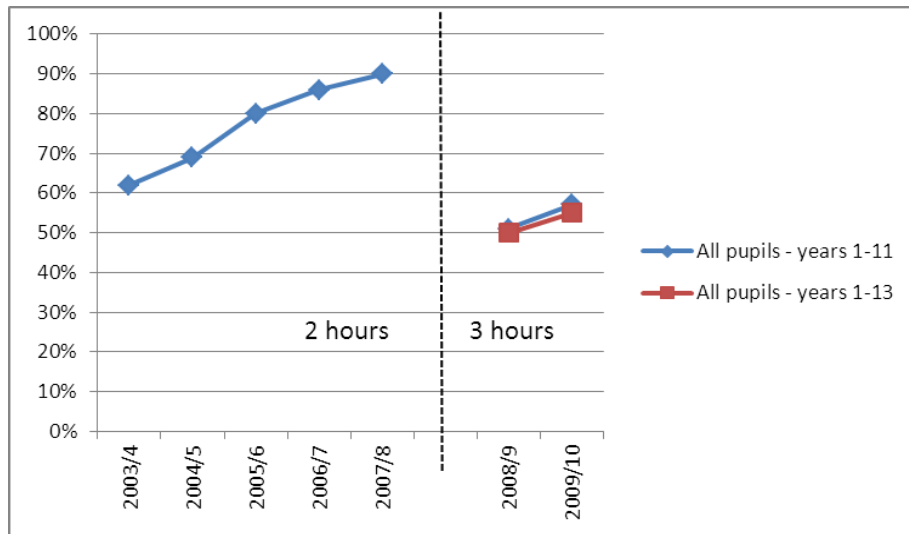
*Recording / Reporting of Data by Agencies / Individuals responsible for Delivery of the Programmes or Projects.*

The first factor which may explain this phenomenon is that the responsibility for recording and reporting local programme / project data was that in many instances, these data were provided by those responsible for project or programme delivery with obvious implications for bias in collection and / or reporting of data. A significant example of this was in the Department for Education annual surveys (2000-2010) of sport participation in which the reporting of the number of children experiencing two (and subsequently three) hours of high quality sport and physical education was undertaken by in-school staff, typically either physical education staff or head teachers, based on data collected by Partnership Managers in schools (see Figure 2). Both the **validity** of the definition of what constitutes 'high quality sport and physical education' and the **reliability** of the calculation of the amount of time / numbers of children involved was thus problematic. It is interesting to compare the results of a survey externally administered by Ipsos Mori (Fraser & Ziff, 2009) whose data suggests a reduced estimate of participation by school age children in sport, though the measures of participation used are slightly different from those used in the PESSYP survey.

In many instances there was no viable alternative to the collection and reporting of data by project / programme providers, and in such cases external audit and verification of the robustness of data collection and reporting was not undertaken.



**Figure 2: Percentage of pupils who participated in at least two or three hours of high quality PE and out of hours school sport in a typical week**



Source: PE and Sport Surveys, Department for Education 2003/4-2009/10

### 3.2.2. The Quality of the Evaluation

Perhaps the most significant limitation in terms of the quality of project and programme evaluations was the failure to address issues of additionality. Additionality as a concept relates to assessing what the overall impact of a project / programme is, by assessing the outcome of an intervention or project (the gross impact) minus any outcomes that would have happened anyway (thus establishing the net impact) enabling us in our case to conclude the size of the net increase in participation brought about by projects / programmes under review.

Here we were interested in establishing what the impacts were of projects / programmes introduced as a product (or by-product) of staging the 2012 Games in London. To do this we needed to take account of the *policy counterfactual* (what would have been provided if the games had not been staged in London) and the *outcome counterfactual* (how many people would have been participating in sport if the Games had not been hosted by London). In arriving at an estimate of the net impact of the intervention, account needs to be taken of *leakage* (for example issues relating to whether the intervention is meeting the needs of the intended constituency or meeting other demands, for example services intended for those with disability being used by the able-bodied population); *substitution* (whether rather than additional sporting activity taking place, participants are substituting taking part in 2012 projects for other non-2012 related projects); *displacement* (when providing new projects displaces existing provision, or when new users of a service displace existing users); and in a positive sense *multiplier effects* (where, for example, participation by those attracted results in a spin-off effect of increased participation by others).

The calculation of the multiplier effect is thus expressed in the following equation in Table 1:

**Table 1: Net Impact Calculation**

<b>Net Impact</b>	=	<b>Additionality of Intervention</b>
<b>Net Impact*</b>	=	[Gross Impact x (1-Leakage) x (1-Displacement) x (1-Substitution) x (1+Multiplier effect)]
*after taking into account of the counterfactuals effects		

Amongst the aspects of additionality which were not pursued in the majority of 2012 project evaluations was the policy counterfactual. For example in the Sportivate project funded as part of the Sport England London 2012 Legacy Programme ‘Places People Play’, additional funding was supplied via the Lottery to support the staging of the programme which involved providing coaching 8-12 sessions followed by an introduction to a club where participation could be sustained (Sport Structures, 2013). However in the evaluation of this initiative the extent of the level of displacement, where for example National Governing Bodies, Local Authorities or sports clubs engage in the provision of opportunities in the Sportivate programme but in doing so may have to limit other forms of non-2012 related provision, was not assessed. We do not know for example the extent to which running the Sportivate schemes displaced other activity previously provided by local authorities or National Governing Bodies, or new participants displaced others in clubs since coaching, equipment and space resources are finite and carrying capacity may be limited. In addition in a number of projects (such as the PESSYP initiative, there was no baseline data available against which to measure change in participation (Grant Thornton, Ecorys, Loughborough University, & Oxford Economics, 2012).

The only project evaluation to incorporate a full account of additionality elements was the PricewaterhouseCoopers report of the *Free Swimming* scheme introduced in England in 2009 (PricewaterhouseCoopers, 2010a, 2010b) aimed at increasing participation among the over-60 and under 16 age groups. The publication of the report ironically led to government’s early abandonment of the scheme because of issues of substitution (existing swimmers reduced attendance at swimming sessions where payment was required, taking advantage of free swims); displacement (some swimmers were crowded out by the free swimmers); and leakage (the target population was the physically less active, but in effect those who already swam, participated more often as a result of this project). In the light of these findings government concluded that the scheme did not provide value for money.

### **3.2.3. The Lack of Evaluation**

In some cases evaluations had not been undertaken and reporting tended to be limited to attendance or user figures. This was the case, for example, for the Gold Challenge which ran from 2010-12. In addition some evaluations which were being undertaken (e.g. on the School Games) had yet to report at the time of undertaking the final meta-evaluation exercise, and it was unclear how such evaluations would take account of additionality issues. In some cases where evaluation or at least monitoring of participation in programmes had been undertaken some were discontinued. For example the ending of collection of data in relation the number of hours participation in quality sport and physical education which had accompanied the PESSYP programme meant that the impact of the Coalition government’s ending of School Sport Partnership funding announced in 2010 as part of a reduction of public spending in response to the recession, could not be assessed

#### **4. Realist evaluation and the identification of what causes achievement of policy outcomes.**

However while evaluations might identify inputs, throughputs, outputs and outcomes, nevertheless an understanding of the factors *causing* particular outcomes in given contexts is neglected. Realist evaluation promoted as an evaluation approach by Pawson and his colleagues (Pawson, 2001; Pawson, 2013; Pawson & Tilley, 1997) seeks to expose the links between the context, causal mechanisms and outcomes in policy programmes or interventions.

Pawson *et al* seek, amongst other things, to distinguish policy evaluation research from the experimental research paradigm in which the RCT is the dominant mode of enquiry. The RCT takes a group exhibiting a certain condition, randomly assigns part of the group to a treatment group, the remainder to a control group which is not subject to the treatment and then seeks to compare the outcome of an intervention by comparing post-intervention the difference between the two groups. Any difference is attributed to the effect of the intervention. This model of evaluation is effectively a closed system in which the relevant variables are known and controlled by the evaluators<sup>3</sup>. Policy evaluation in social contexts is by contrast subject to enormous variation. Thus while medical research operating with the RCT seeks to answer the question of ‘what works?’ by (in theory) isolating the impact of a single variable, policy evaluation addresses a much more wide ranging set of questions, namely ‘what works? in what ways? for which types of group? under which context or conditions?’ because such evaluation takes place within open systems in which account has to be taken of a far wider set of variables.

Realist evaluation starts with a programme theory of change. This hypothesises the explicit or assumed link between the inputs into a policy system, the throughputs (or activities or interventions undertaken), the immediate outputs, and the longer term outcomes or impacts of the programme. It then subjects this set of assumed causal links to rational and empirical tests to support or undermine causal explanations. Whereas comparison of intended and actual outcomes may provide an account of whether, and the degree to which, a policy has met with success, Realist Evaluation seeks to uncover explanations of why it has or has not succeeded. In other words we may be able to say whether or not the staging of the Games is associated with an increase in participation in sport and physical activity, but without testing a programme theory of change, we will not necessarily be able to say why. In such circumstances, without context-specific causal explanations, it may be difficult to emulate policy success or avoid failure.

While realist evaluation has gained ground in recent years in a wide range of policy contexts, its application in the sports field remains relatively neglected. We cite here a recent study however (Chen & Henry, 2015), undertaken within the context of an evaluation of the impact of the 2012 Games on a non-hosting region which seeks to drill down to the implicit and explicit theory(ies) of change underpinning particular local policies. In this study, the authors unpack through document analysis of strategy documents and operational plans, and qualitative data from interviews with key stakeholders, to understand exactly how a particular policy initiative has been designed to produce an increase in sport and physical activity within selected work organisations.

---

<sup>3</sup> In fact the concept of closed system is an oversimplification. The ability to control all variables at play in a given RCT is likely to be far from perfect and thus the system is perhaps better described as being treated as relatively closed.

The initiative subject to evaluation in the paper by Chen and Henry was a programme called the Work Place Challenge which was a free, online competition between businesses running in 2011 and 2012 that allowed participants to log their level of physical activity on a shared database over the course of the programme. The programme was organised and administered by the Leicester Shire and Rutland County Sport Partnership, and its aim was to stimulate competition between work organisations in terms of the recorded levels of sport and exercise undertaken by employees over a given period. Prizes were offered to encourage continued participation in WCP for individuals as well as for the organisation or workplace. A range of toolkits, resources and forms of support were made available by the Leicestershire County Sport Partnership, to help the workplace organisation to actively engage with the programme. Chen and Henry’s evaluation study of the WCP also involved collecting data from two waves of questionnaires in 2011 (n=125) and 2012 (n=77) and qualitative data from interviews with participants and programme deliverers (n=15)<sup>4</sup>.

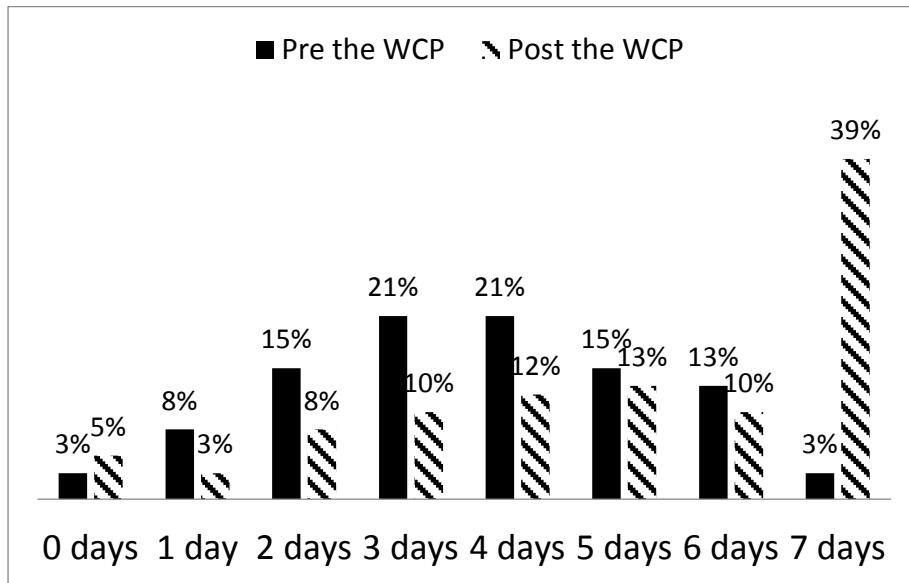
The logic which underpinned the WCP initiative was uncovered in documentary analysis and interviews with stakeholders and is summarised in Table 2 below, while the level of participation before and after the staging of the WCP competition in 2012 is reported in Figure 3, highlighting significant increases in sport and exercise participation.

**Table 2: Summary of the Elements of the Logic Underpinning the Workplace Challenge ‘Programme Theory’.**

<ol style="list-style-type: none"> <li>1. The publicity accorded to the staging of the Olympics in the UK provides a focus for publicity concerning sport</li> <li>2. The affective impact (for many people) of the prospect of the proximity of this megaevent generates greater interest in, and a positive response to, sport-related lifestyles for a significant proportion of the population</li> <li>3. The Leicestershire Steering Group/Inspire Leicestershire provided information on the benefits of exercise, the required levels of participation and intensity required to generate health benefits, as well as information on how participation in sport and activity in the workplace, or among the workforce, can be increased</li> <li>4. The incentive to participate in sport and physical activity will be strengthened, and tendencies for recidivism will be undermined, if records of progress are kept and performance, in terms of maintaining increased levels of participation, rewarded.</li> <li>5. The provision of opportunities for sport (in intra- and inter-organisational competitions) will provide opportunities for social contexts (e.g. teamwork promoting social bonding), which have the potential to make exercise and the context of exercise more enjoyable.</li> <li>6. Competition between organisations in terms of levels of exercise undertaken by the workforce, and measurement and recording of exercise levels, will motivate employees to sustain and improve their performance</li> <li>7. The increase in physical activity that is thus promoted will enhance the physical health of those members of the workforce who participate, and thus will reduce public health COSTS</li> </ol>
--

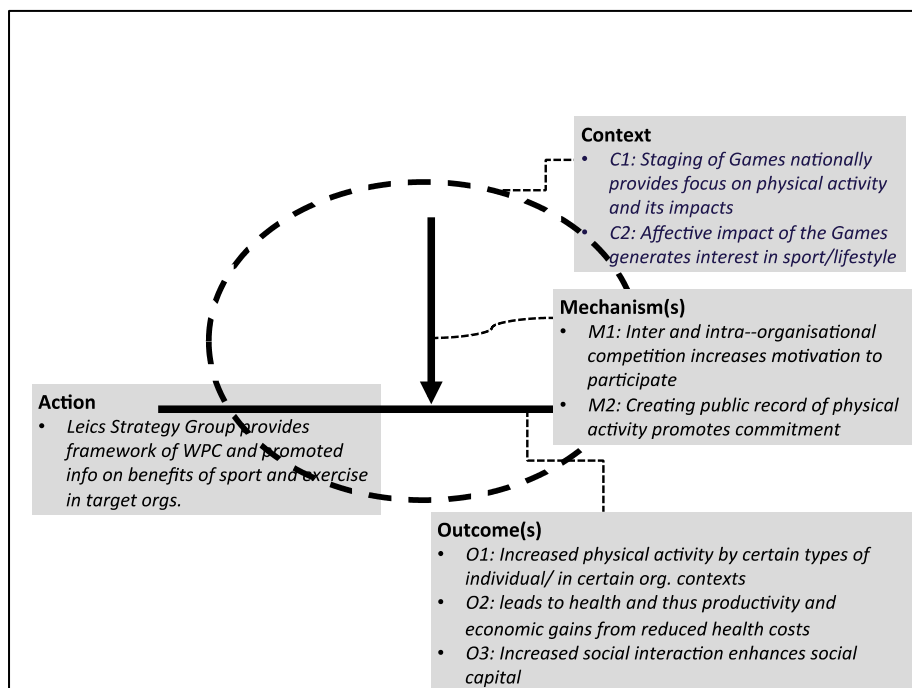
<sup>4</sup> One should note that since the samples for the evaluation were the result of self selection (people opted to provide data), the sample is likely to have been skewed towards those who may have been positively predisposed to the programme.

**Figure 3: Comparison of the days of moderate intensity physical activity (at least 30 minutes) per week before and after the 2012 Workplace Challenge programme**



However the data reported in Figure 2 gives a relatively crude picture of whether and how the intervention has been successful. Chen and Henry in pursuing a realist evaluation agenda go on to highlight not only the causal mechanisms (and / or barriers) which are evident (e.g. effectiveness of intrinsic and extrinsic motivators; positive and negative organisational cultures) but also to define for whom those causal mechanisms work (e.g. gender; age; previous sporting / exercise history); in what ways the mechanisms work (e.g. promoting sport or non-sport exercise competitive or recreational sporting activity; episodic or regular participation); in which kinds of circumstances or context (e.g. in different kinds of employing organisation, in organisations where there is a recognized organisational champion of the WCP, whether work is physically active or sedentary); and the kinds of outcome the causal mechanisms promote (e.g. the types of, and the regularity and sustainability of, exercise). An abbreviated form of the types of explanation developed is provided in Figure 4.

**Figure 4: An Abbreviated Illustration of the Development of a Realist Explanation of the Workplace Challenge**



Thus the realist evaluation approach represents a significant advance on approaches to evaluation in that it provides contingent, and context-specific explanations of what causes outcomes to be achieved

#### 4. Conclusions

In this paper we have sought to outline the application of the meta-evaluation approach to assessing aspects of the legacies of the 2012 London Olympic and Paralympic Games. The quality of meta-synthesis of data from qualitative and quantitative evaluations, a key element in meta-evaluation, is necessarily dependent on the quality and rigour of the constituent evaluations the results of which are to be synthesised. In our review of the approach as applied to an evaluation of the legacies of the 2012 Games, we highlight validity and reliability problems associated with the data collected in constituent evaluations; a lack of evaluation per se in constituent programmes / projects (as opposed to simple monitoring data); and the absence of consideration of additional factors which limit calculation of the net effect of staging the Games. These difficulties may be summed up in the claim that full synthesis of evaluations is only possible if certain quality thresholds are respected in each of the evaluations included within the synthesis.

Finally, identifying not simply what has changed as a result of staging the Games but why it has changed – in other words providing a causal account of change – is critical if we are to understand the contingent and context-specific factors explaining policy success or failure. Realist evaluation offers the prospect of such explanation, and the example of evaluation of the Workplace Challenge programme discussed in this article seeks to point to the possibility of more robust and fruitful approaches to the assessment of the nature of policy impacts and the causal mechanisms underlying such impacts.

## References

- Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: a critical review. *Bmc Medical Research Methodology*, 9(59).  
<http://www.biomedcentral.com/content/pdf/1471-2288-9-59.pdf>  
doi:10.1186/1471-2288-9-59
- Chen, S., & Henry, I. (2015). Evaluating the London 2012 Games' impact on sport participation in a non-hosting region: a practical application of realist evaluation. *Leisure Studies*. doi:10.1080/02614367.2015.1040827
- Chen, S., Henry, I., & Ko, L.-M. (2013). Meta-evaluation, analytic logic models and the assessment of impacts of sports policies. In I. Henry & L.-M. Ko (Eds.), *Handbook of Sport Policy*. London: Routledge.
- Crombie, I. K., & Davies, H. T. O. (2009). What is Meta-analysis? *What is ... ? series*.  
<http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/meta-an.pdf>  
Retrieved from  
<http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/meta-an.pdf>
- Department of Culture Media and Sport. (2008). *Before, during and after: making the most of the London 2012 Games*. London: Department of Culture Media and Sport.
- Department of Media Culture and Sport. (2008). *London 2012 Olympic and Paralympic Games Impacts and Legacy Evaluation Framework*. Retrieved from London:  
[http://www.gamesmonitor.org.uk/files/DCMS\\_Olympic\\_Evaluation\\_final\\_report.pdf](http://www.gamesmonitor.org.uk/files/DCMS_Olympic_Evaluation_final_report.pdf)
- Department of Culture Media and Sport. (2010). *Plans for the legacy from the 2012 Olympic and Paralympic Games*. Retrieved from
- Department of Culture Media and Sport. (2012). *Taking Part 2012/13 Quarter 1: statistical release*. Retrieved from London:  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/77537/Taking-Part\\_2012-13\\_Quarter-1\\_Report.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77537/Taking-Part_2012-13_Quarter-1_Report.pdf)
- Featherstone, D., Ince, A., Mackinnon, D., Strauss, K., & Cumbers, A. (2012). Progressive localism and the construction of political alternatives. *Transactions of the Institute of British Geographers*, 37(2), 177-182.
- Fraser, J., & Ziff, A. (2009). *Children and Young People's Participation in Organised Sport, Omnibus Survey*, . Retrieved from London:
- Gough, G., Martin, P., Grant Thornton, & ECORYS. (2012). *Meta-evaluation of the Impact and Legacy of the London 2012 Olympic Games and Paralympic Games: Developing Methods Paper* Retrieved from London:
- Grant Thornton, Ecorys, Loughborough University, & Oxford Economics. (2012). *Report 4: interim evaluation*. Retrieved from London:  
<https://www.gov.uk/government/publications/report-4-interim-evaluation-meta-evaluation-of-the-impacts-and-legacy-of-the-london-2012-olympic-games-and-paralympic-games>
- Greenhalgh, T., Wong, G., Westhorp, G., & Pawson, R. (2011). Protocol - realist and meta-narrative evidence synthesis: Evolving Standards (RAMESES). *Bmc Medical Research Methodology*, 11. doi:10.1186/1471-2288-11-115
- House of Lords. (2013). Keeping the flame alive: the Olympic and Paralympic Legacy. *Select Committee on Olympic and Paralympic Legacy - First Report* Retrieved from  
<http://www.publications.parliament.uk/pa/ld201314/ldselect/ldolympic/78/7802.htm>
- Lee, M. (2006). *The Race for the 2012 Olympics*. London: Virgin Books.
- Lucas, P., Arai, L., Baird, R., Law, C., & Roberts, H. M. (2007). Worked examples of alternative methods for the synthesis of qualitative and quantitative research in systematic reviews. *BMC Med Res*

- Meth 2007, 7(4):. *Bmc Medical Research Methodology*, 4.  
<http://www.biomedcentral.com/content/pdf/1471-2288-7-4.pdf>  
 doi:doi:10.1186/1471-2288-7-4
- Noblit, G. W., & Hare, R. D. (1988). *MetaEthnography: synthesising qualitative studies*. Newbury Park: Sage.
- Pawson, R. (2001). *Evidence based policy : II. the promise of 'realist synthesis'*: ESRC UK Centre for Evidence Based Policy and Practice.
- Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. Leeds: University of Leeds.
- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.
- Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide* Wiley-Blackwell (21 Oct 2005).
- PricewaterhouseCoopers. (2010a). *Evaluation of the impact of Free Swimming: Year 1 report – main report*. Retrieved from London:
- PricewaterhouseCoopers. (2010b). *Evaluation of the impact of Free Swimming: Year 1 report annexes*. Retrieved from London:  
[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/77393/FSPyear1\\_annexes.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77393/FSPyear1_annexes.pdf)
- Ray, L. (2012). Community Engagement and the Olympics. *Journal of the Society of Archivists*, 33(1), 1-8. doi:10.1080/00379816.2012.666170
- Scriven, M. (1969). An introduction to meta-evaluation. *Educational Products Report*, 2(1), 36-38.
- SecEd. (2013). Ofsted tells DfE it must build on success of scrapped School Sport Partnerships. *SecEd*, 21 February 2013. <http://www.sec-ed.co.uk/news/ofsted-tells-dfe-it-must-build-on-success-of-scrapped-school-sport-partnerships> Retrieved from <http://www.sec-ed.co.uk/news/ofsted-tells-dfe-it-must-build-on-success-of-scrapped-school-sport-partnerships>
- Sport England. (2015a). The Active People Survey. Retrieved from <http://www.sportengland.org/research/about-our-research/what-is-the-active-people-survey/>
- Sport England. (2015b). Once a week participation in sport (1 x 30 minutes moderate intensity). *Active People Survey Active People Survey 9 - Q2 April 2014 – March 2015*. Retrieved from [http://www.sportengland.org/media/875700/1x30\\_overall\\_factsheet\\_aps9q2v2.pdf](http://www.sportengland.org/media/875700/1x30_overall_factsheet_aps9q2v2.pdf)
- Sport Structures. (2013). *Sportivate Programme Evaluation: Year two annual report April 2012 – March 2013*. Retrieved from London:
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *Bmc Medical Research Methodology*, 8(45).
- UK Government. (2015). *2010 to 2015 government policy: 2012 Olympic and Paralympic legacy*. Retrieved from London:  
<https://www.gov.uk/government/publications/2010-to-2015-government-policy-2012-olympic-and-paralympic-legacy/2010-to-2015-government-policy-2012-olympic-and-paralympic-legacy>
- UK Government, & Mayor of London. (2013). *Inspired by 2012: The legacy from the London 2012 Olympic and Paralympic Games - A joint UK Government and Mayor of London report*. Retrieved from London: