

Northumbria Research Link

Citation: Kotecha, Birju (2017) The ICC's Office of the Prosecutor and the Limits of Performance Indicators. *Journal of International Criminal Justice*, 15 (3). pp. 543-565. ISSN 1478-1387

Published by: Oxford University Press

URL: <https://doi.org/10.1093/jicj/mqx028> <<https://doi.org/10.1093/jicj/mqx028>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/30965/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

The ICC's Office of the Prosecutor and the Limits of Performance Indicators

Birju Kotecha*

Abstract

Performance indicators have long been part of the vocabulary of businesses and public administrations. Essentially, they are market-oriented tools occupying a central role in both the management and measurement of organizational performance. The present article assesses, in particular, the performance indicators adopted by the International Criminal Court's (ICC) Office of the Prosecutor (OTP). In so doing, it traces the development of such indicators within the OTP strategies, and argues that current indicators do not measure prosecution effectiveness in terms of outcomes within conflict-affected societies. The article corroborates this argument by explaining how current indicators are inclined to prioritize the economy and efficiency of prosecutions, thus producing a gap in the measurement of the prosecutions' effectiveness. The mentioned gap is explained by the OTP's fundamental lack of clarity in precisely identifying the outcomes of prosecutions. In turn, the performance indicators adopted by the OTP may engender perverse effects, further aggravating the lack of clarity concerning outcomes. The article recommends the adoption of principles aimed at improving future OTP performance measurement.

1. Introduction

Performance is an ambiguous and contested concept. On one hand, it denotes capability and success, and often exudes 'an aroma of action, dynamism and purposeful effort'.¹ However, in a common and often pejorative sense, performance can mean an overstated demonstration or presentation.² This ambiguity is most common within businesses and organizations, where performance is juxtaposed to two further concepts: management and measurement. The relationship is aptly surmised: 'to manage performance it is also essential to measure performance'.³ Hence, 'everyone is measuring performance'.⁴ The motivation for performance measurement is based on the oft-cited beliefs that 'what gets measured gets done'⁵ and 'if you don't measure results, how can you tell success from failure'.⁶ Once a commitment to measurement has been made, the only remaining question is how to measure

* I would like to thank David McGrogan, Robert C. Blitt and Sarah Dadush for helpful comments on a draft of this article. I am also grateful to Luke Moffett and Alison Bisset for feedback on a much earlier version. [birju.kotecha@northumbria.ac.uk]

¹ C. Pollit, 'Beyond the Managerial Model: The Case for Broadening Performance Assessment in the Government and the Public Services', 2 *Financial Accountability and Management* (1986) 155-169, at 160.

² A. Stevenson (ed.), *Oxford Dictionary of English* (3rd edn., Oxford University Press, 2010), at 1320.

³ International Criminal Court Office of the Prosecutor (ICC OTP), *Strategic plan June 2012-2015*, 11 October 2013, § 82.

⁴ R.D Behn, 'Why Measure Performance? Different Purposes Requires Different Measures', 63 *Public Administration Review* (2003) 586-606, at 586

⁵ D. Osborne and T. Gaebler, *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector* (Addison-Wesley, 1992), at 146.

⁶ B. Townley, D.J. Cooper and L. Oakes, 'Performance Measures and the Rationalisation of Organizations', 24 *Organization Studies* (2003) 1045-1071, at 1060.

performance. A measurement tool first made fashionable during public administration reforms in the 1980s,⁷ performance indicators provide standards that can help describe progress towards established goals, and thus are central to measurement practice.⁸

From hospital waiting list times, to graduation numbers, to user satisfaction scores, to rates of recorded crime, there is virtually no area of public life where performance indicators do not extend. Their significance and prevalence has been transformative. The significance of indicators is based on their perceived utility. They can focus action plans, help budget, motivate workforces and provide feedback to help improve performance.⁹ Their prevalence reflects the importance attached to holding organizations accountable, and thus indicators are central to what has been described as an ‘audit explosion.’¹⁰ Accordingly, performance indicators speak to many audiences beyond those concerned with day-to-day management. These include funding and supervisory bodies concerned with audit, monitoring and oversight, and to whom an organization is accountable.¹¹

In this context, this article assesses the performance indicators of the International Criminal Court’s (ICC) Office of the Prosecutor (OTP). To date, literature on the Court has mostly overlooked performance indicators. For instance, there is little scrutiny on the Court-wide attempt to develop such indicators.¹² This lack of research is surprising, considering that the adoption of performance indicators is a novel feature of the ICC, with no comparison at other international criminal tribunals. Similarly, there is no literature on the OTP’s development of performance indicators, in particular the set of indicators identified by its

⁷ N. Carter, ‘Performance Indicators: ‘back-seat driving or ‘hands off control’, 17 *Policy and Politics* (1989) 131-138, at 132.

⁸ N. Carter, ‘Learning to Measure Performance: The Use of Indicators in Organizations’, 69 *Public Administration* (1991) 85-101, at 90. See also K.E. Davis, B. Kingsbury, S.E. Merry, ‘Introduction: Global Governance by Indicators’, in K.E. Davis et al. (eds), *Governance by Indicators: Global Power Through Quantification and Rankings* (Oxford University Press, 2012) 3-28.

⁹ Behn, *supra* note 4, at 588.

¹⁰ M. Power, *The Audit Society: Rituals of Verification* (Oxford University Press, 1997), at 6-10.

¹¹ Pollit, *supra* note 1, at 163.

¹² ICC Assembly of States Parties (ASP), *Strengthening the International Criminal Court and Assembly of States Parties*, ICC-ASP/13/Res.5, 17 December 2014, § 7(b) calling on ‘the Court to intensify its efforts to enhance the efficiency and effectiveness of proceedings [and] to intensify its efforts to develop qualitative and quantitative indicators that would allow the Court to demonstrate better its achievements and needs, as well as allowing State Parties to assess the Court’s performance in a more strategic manner.’ To date, the only non-ICC literature on ICC performance indicators is constituted by briefings and consultation documents that have been developed in collaboration with the Court. See Open Society Foundations Open Society Justice Initiative, *Establishing Performance Indicators for the International Criminal Court Briefing Paper*, November 2015, available online at <https://www.opensocietyfoundations.org/sites/default/files/briefing-icc-perforamnce-indicators-20151208.pdf> (visited 1 June 2017); Federal Department of Foreign Affairs of Switzerland, ‘International Criminal Court Retreat on Performance Indicators’, 24 May 2016, available online at https://www.eda.admin.ch/content/dam/eda/en/documents/aussenpolitik/voelkerrecht/2016-criminal-court-conveno-srummy_EN.pdf (visited 1 June 2017).

current prosecutorial strategy.¹³ This latter research gap is even more surprising and significant, for two reasons.

First, the choice of certain performance indicators (as opposed to others) embodies a normative claim about the appropriate standards for evaluating the conduct of actors.¹⁴ As Davis, Kingsbury and Merry contend, indicators are embedded with a model for an ideal conduct and governance and, thus, the best process for achieving it.¹⁵ Hence, the OTP's selection of indicators expresses a policy and a vision of what is desirable, and thus explains what the Office interprets as excellent performance.¹⁶

Second, performance indicators have a communicative element that resonates with diverse audiences within political and social environments.¹⁷ These indicators can communicate with and seek to influence the expectations of external constituencies. With regard to the OTP, it is worth noting that the Office's conduct is often perceived as synonymous with the Court's conduct. The Prosecutor is the ICC's most public-facing actor, and the OTP's functions are critical for the Court's operation. As such, it hardly needs to be said that the OTP's activities, e.g. the selection of situations and cases, are the most common target for criticism about the Court. Hence, indicators are also an institutional response to the conflict-affected and political environments within which the OTP's decisions are perceived and inevitably contested. Perhaps more importantly, these indicators also speak to the Office's attempt at having an ultimate impact within these environments. Thus, performance indicators reveal how the Office intends to contribute — and how it indeed contributes — to the Court's goals of ending impunity and preventing crimes, alongside a number of often-labelled idealistic goals, such as peace, stability and reconciliation.¹⁸

The article conducts its assessment of performance indicators with the use of public administration perspectives, positioning itself at a disciplinary intersection that is described as 'international bureaucracy research'.¹⁹ At the heart of this disciplinary intersection lies the recognition that, historically, international organizations have adopted management

¹³ See ICC OTP, *Strategic plan June 2012-2015*, 11 October 2013, §§ 94-98.

¹⁴ Davis, Kingsbury and Merry, *supra* note 8, at 9-10.

¹⁵ *Ibid.*

¹⁶ W.N. Espeland and M. Sauder, 'The Dynamism of Indicators', in Davis et al. (eds), *supra* note 8, 86-109, at 87, 93.

¹⁷ Davis, Kingsbury and Merry, *supra* note 8, at 10.

¹⁸ For one of many sceptical accounts, see P. Akhavan, 'The Rise and Fall and Rise of International Criminal Justice', 11 *Journal of International Criminal Justice* (2013) 527-536.

¹⁹ J. Ege and M.W. Bauer, 'International bureaucracies from a Public Administration and International Relations perspectives', in B. Reinalda (ed.), *Routledge Handbook of International Organization* (Routledge, 2013) 135-149, at 135.

approaches that progressively mirrored national public management reforms.²⁰ This disciplinary perspective also provides insight into international organizations' strategy-making and, in turn, points to the context of performance indicators.²¹ Indeed, traditionally, public administration literature does not regard indicators as an isolated management tool, but rather as one component of a suite of market-oriented principles known as 'managerialism'.²²

Performance indicators (often prefixed with the qualifier 'key') sprung originally within highly managerial environments. Within these environments, indicators became routinely associated with maximising the virtuous three 'E's of performance: economy, efficiency and effectiveness.²³ These three performance concerns incorporate the relationship between organizational inputs, outputs and outcomes. In short, economy is concerned with performance that minimises inputs (most often understood in terms of financial costs and resources). In light of those costs, efficiency targets performance that maximises outputs – i.e. the day-to-day service activities that an organization delivers or provides (e.g., schools deliver educational qualifications and hospitals provide treatments and operations — these are their outputs). Finally, effectiveness is concerned with performance that transforms successful outputs into particular outcomes within target audiences.²⁴ As such, outcomes are concerned with real-world phenomena occurring outside the organization, and these outcomes, in turn, are intended to support the original and fundamental goals set by an organization. For example, high quality school qualifications can achieve specific outcomes (e.g. to make students literate and numerate, or to develop their critical reasoning skills, etc.) that produce well-informed and skilled citizens, or successful hospital treatments and operations can achieve specific outcomes (e.g. to make patients free from ill-health

²⁰ E. Missoni and D. Alesani (eds), *Management of International Institutions and NGO's: Framework, Practices and Challenges* (Routledge, 2013), at 3.

²¹ Ege and Bauer, *supra* note 19, at 135.

²² Managerialism (often synonymous with 'New Public Management') has been described by Christopher Pollit as a set of beliefs, practices and values that sees better management as the effective solution to economic and social ills. The origins of managerialism are commonly traced to the ideological and market-oriented reforms of traditional Weberian public bureaucracies in the 1980s and 1990s. See C. Pollitt, *Managerialism and the Public Services* (2nd edn., Blackwell Publishing, 1996) 1-10; C. Hood, 'A Public Management for All Seasons?' 69 *Public Administration* (1991) 3-19, at 4-5. For a recent review of managerialism, see C. Pollitt, 'Managerialism Redux?' 32 *Financial Accountability and Management* (2016) 429-446; T. Kilkaer, 'What is Managerialism?' 41 *Critical Sociology* (2015) 1103-1119 and D.G. Mathiasen, 'The New Public Management and its Critics', 2 *International Public Management Journal* (1999) 90-111, at 92-93.

²³ N. Carter, R. Klein and R. Day, *How Organizations Measure Success: The Use of Performance Indicators in Government* (Routledge, 1995), at 35-36.

²⁴ *Ibid.*

symptoms, physically mobile, or to help them self-manage pain etc.) that lead to long-term health and well-being.²⁵

Thus, in the case of the OTP, performance can be translated into managing its running costs (*economy/inputs*), into maximizing prosecution activities in light of those costs (*efficiency/outputs*) and, finally, into ensuring that successful prosecutions can achieve outcomes that further the goals of the Rome Statute, i.e. ending impunity and contributing to the prevention of crimes (*effectiveness/outcomes*).

This article begins by tracing the historical development of performance indicators within the Office's prosecutorial strategy. Then, it adopts the mentioned three-'E's framework in order to assess OTP's performance indicators. This framework corroborates the article's main contention that the OTP's indicators do not sufficiently measure the effectiveness of prosecutions in terms of their intended outcomes, particularly within conflict-affected societies. In this light, the article concludes by offering principles that can improve the future measurement of effectiveness.

2. Prosecutorial Strategy and Performance Indicators

First of all, it is important to consider the OTP's motivation for establishing performance indicators. This move was evidently inspired by a Court-wide commitment to managerialism. Indeed, the ICC's first Strategic Plan explicitly makes one of the Court's goals to be a 'model of public administration.'²⁶ The Plan describes this goal as focused on results rather than processes, empowering managers to take decisions to increase activities' speed, quality and efficiency, and to guarantee accountability within a non-bureaucratic administration.²⁷ This language chimes with characteristic managerial methods that are 'hard-headed, business-minded, cost-conscious, and data-driven'.²⁸

²⁵ C. Pollitt and S. Dan, *The Impacts of the New Public Management in Europe: A Meta-Analysis*, European Commission, Coordinating for Cohesion in the Public Sector of the Future, 14 December 2011, available online at http://www.cocops.eu/wp-content/uploads/2012/03/WP1_Deliverable1_Meta-analysis_Final.pdf (visited 20 May 2017), at 11-12.

²⁶ ICC ASP, *Strategic Plan of the International Criminal Court*, ICC-ASP/5/6 (5th Session), 5 August 2006, at 9-10.

²⁷ *Ibid.* The Strategy's reference to a 'non-bureaucratic' administration alludes to preventing bureaucratic cultures that are disproportionately concerned with procedural correctness, rather than with efficiency and productivity. This priority is underlined in the ICC's current Strategic Plan, where the declared mission of the Court includes the fair and effective conduct of investigations and prosecutions, administrative transparency, efficiency and accountability. See *International Criminal Court Strategic Plan (2013-2017)*, 24 July 2015, at 2.

²⁸ C. Hood and R. Dixon, *A Government that Worked Better and Cost Less? Evaluating Three Decades of Reform and Change in UK Central Government* (Oxford University Press, 2015), at 15.

The goal of being a model of public administration also establishes a ‘One Court’ structure, whereby the OTP works with other Court organs on matters of common concern.²⁹ This framework maintains respect for the OTP’s independence but allows for the OTP’s ‘maximal integration’ into the Court, especially within strategic planning.³⁰ Within this context, the OTP’s first prosecutorial strategy fully subscribed to the Court’s goal to be a model of public administration and committed to devising performance indicators.³¹

Since that moment, the development of performance indicators has been an incremental and decade-long process, evidence of which is found within four OTP’s strategic documents.³²

In the beginning, the OTP’s first strategic document made a cursory statement of intent to ‘establish a clear set of performance indicators and evaluation processes’ while recognizing that — as no trials had yet commenced — it was too early to start evaluating its performance.³³

Four years on, and in spite of trials having commenced, the second strategy affirmed a commitment to participate in the Court-wide work on performance indicators. The strategy recognized that measuring the Office’s performance is a complex task that requires an evaluation of the entire Rome Statute’s system. The strategy highlighted that this evaluation should recognize the role and function of complementarity, and how co-operation is fundamental for the system’s effective functioning.³⁴ The strategy also provided, rather tentatively, examples of ‘diversified performance indicators’.³⁵ These included measuring the number of national prosecutions for international crimes, of peace accords that exclude amnesty for ICC crimes, of public declarations of political leaders in support of the Court’s decisions, and the reduction in the number of armed groups and armed forces using children.³⁶ Nonetheless, these were solely included for the purpose of demonstrating what future indicators may look like.

One might have thought that by the time of the third strategy a concrete set of relevant performance indicators would have been developed. By contrast, the third strategy cited

²⁹ ICC ASP, *Strategic Plan of the International Criminal Court*, ICC-ASP/5/6 (5th Session), 5 August 2006, at 3-4

³⁰ *Ibid.* See also Art. 42(1) ICCSt. (The Office of the Prosecutor).

³¹ ICC OTP, *Report on Prosecutorial Strategy*, 14 September 2006, §§ 9-16.

³² The first was in the form of a report on prosecutorial strategy, and the second merely a prosecutorial strategy. Both the third and current strategy are in the form of strategic plans. For simplicity, the term ‘strategy’ is adopted in this article.

³³ ICC OTP, *Report on Prosecutorial Strategy*, 14 September 2006, §§ 13-16.

³⁴ ICC OTP, *Prosecutorial Strategy 2009-2012*, 1 February 2010, § 79.

³⁵ *Ibid.*, §§ 78-80.

³⁶ *Ibid.*

measurement complexity and a lack of time and resources to explain the continuing absence of clear performance indicators.³⁷ According to the strategy, measurement complexity was due, on one side, to the challenge of determining the precise outcome of OTP’s activities (offering the prevention of crimes as an example of such an outcome) and – on the other side, to accurately measuring the OTP’s impact on outcomes, which may have been determined – alternatively or cumulatively – by other causes. Hence, the strategy retreated from measuring outcomes and declared an intent to (re)evaluate whether a ‘refocused and limited’ set of indicators would be more suitable.³⁸ To underline the retreat, the strategy hinted at what such indicators could be, making a reference to ‘securing convictions ... and [to] an increase of the percentage of charges confirmed and [of the] conviction rate.’³⁹

The first formal performance indicators surfaced when the current strategy was released, in July 2015. Again, the strategy grapples with contributory causes and measurement complexity. It explains that the OTP does not have sufficient control over certain outcomes, for instance on crime prevention.⁴⁰ However, the strategy identifies a start-up of fourteen interrelated indicators, reproduced below.⁴¹

Area of Performance	Indicator
Strategic	
Effectiveness	<ol style="list-style-type: none"> 1. Prosecutorial results in terms [of] perpetrators <ul style="list-style-type: none"> • Arrest warrants’, summonses to appear granted / requested • Persons confirmed / charged • Persons convicted / charged 2. Counts granted/counts charged at stage or arrest warrant, confirmation, trial
Operational	
Operational Excellence	
<i>Quality</i>	<ol style="list-style-type: none"> 3. Pattern of judicial findings on how the Office conducts its preliminary examinations, investigations and prosecutions 4. Compliance with Key Office Policies and Standards 5. Quality of interaction with the Office
<i>Efficiency</i>	<ol style="list-style-type: none"> 6. Yearly achieved efficiency gains
<i>Productivity</i>	<ol style="list-style-type: none"> 7. Milestones per activity: planned versus actual.

³⁷ ICC OTP, *Strategic plan June 2012-2015*, 11 October 2013, §§ 94-96.

³⁸ *Ibid.*, § 95.

³⁹ *Ibid.*

⁴⁰ ICC OTP, *Strategic Plan | 2016-2018*, 6 July 2015, § 104.

⁴¹ *Ibid.*, §§ 104-106.

Management Excellence <i>HR</i> <i>Financial Management</i> <i>Risks</i>	8. Implementation of training program per year: planned versus actual 9. Working climate survey 10. Evolution of fitness of work 11. Evolution of the general gender and nationality balance per year 12. Yearly implementation rate of the budget 13. Implementation rate of measures to control priority risks out of the OTP risk register
Innovation and Learning	14. Impact of improvement projects on effectiveness, operational excellence and management excellence.

As displayed, the fourteen indicators are divided in two categories: strategic and operational. The strategic category includes those indicators intended to measure the Office’s effectiveness in achieving its mandate under the Rome Statute.⁴² Operational indicators measure the implementation of the OTP strategy, which is assumed to impact positively on the strategic indicators.⁴³ Hence, the twelve operational indicators are intended to foster the achievement of the two pivotal strategic performance indicators.

As shown in the table, the first strategic indicator — labelled ‘prosecutorial results in terms [of] perpetrators’ — is concerned: with the number of arrest warrants or summonses to appear granted, out of those requested; with the number of defendants whose charges have been confirmed, out of those for whom charges were presented; and the number of persons who have been ultimately convicted, out of those whom were charged. The second indicator refers to the number of counts granted, out of those charged at the stage of the arrest warrant, confirmation or trial.⁴⁴ As the annex of the strategy makes clear, these indicators are intended to meet the strategic goal of conducting impartial, independent, high quality preliminary examinations, investigations and prosecutions.⁴⁵ Thus both the strategic indicators are primarily concerned with quantitative and qualitative outputs, intending to motivate prosecutorial ‘wins’ and dis-incentivise the construction and presentation of requests, charges and cases unlikely to succeed.

Strikingly, the table of indicators also reveals a different interpretation of the three E’s of performance. For instance, both strategic indicators are juxtaposed with effectiveness, and thus, in contrast to the understanding outlined at the outset of the article, the OTP interprets

⁴² *Ibid.*, § 105.

⁴³ *Ibid.*, §§ 105-106.

⁴⁴ *Ibid.*, at § 106.

⁴⁵ *Ibid.*, at Annex 4.

effective performance solely in terms of the success of its output activities — evidenced in positive rulings and verdicts. There is no recognition that effectiveness is normally measured against the real-world outcomes produced upon the success of outputs, or the relationship between those outcomes and the ultimate goals of the Court. Similarly, the sixth operational performance indicator refers to efficiency gains but its detachment from outputs suggests that the OTP understands efficiency more in terms of economy, rather than *a priori* maximising successful outputs with the least amount of resources.⁴⁶

Leaving these concerns aside, what is left is a set of performance indicators that are exclusively output-focused and, as such, they invite general concerns about their limitations. Typically, these indicators reduce what is a complex phenomenon into raw data.⁴⁷ This conversion can be attractive because it standardizes information — making data easy to read and to compare — but, in doing so, it strips meaning and context from the particular phenomena the indicators seek to capture.⁴⁸ In this sense, output indicators always act like ‘tin-openers’.⁴⁹ They do not provide authoritative conclusions about performance, but rather prompt further qualitative inquiry because the performance ‘reading’ is contestable. At the OTP, this can be illustrated by two examples.

To take the first, prosecution results do not account for diverse causal factors that may contribute to their achievement rates. The OTP has recognized this, and has supplemented its performance indicators by creating ‘external critical success factors’. These include the presence or lack of co-operation (i.e. the general support received by the Office) as well as security (i.e. the inability to conduct operations due to security concerns and the impact of witness interference).⁵⁰ Presumably, the reading of prosecution results indicators will need to be reconciled with an accompanying assessment of these critical success factors.

The second illustration concerns the common risk of skewed data contributing to those results. Current OTP strategy provides an example of this risk with its statistics on the rate of charges confirmed at the confirmation stage.⁵¹ The OTP admits (conveniently in a footnote) that the positive performance increase from its previous strategy is rather drastically

⁴⁶ Annex 2 of the current OTP strategy provides a table of efficiency gains. This table summarizes a range of cost savings that are not directly concerned with prosecution outputs — from the preference for economy class flights, to the use of residential accommodation (rather than hotels) during field trips and to savings made by improvements to IT systems.

⁴⁷ Davis, Kingsbury and Merry, *supra* note 8, at 6.

⁴⁸ Davis, Kingsbury and Merry, *supra* note 8, at 8. See also Espeland and Sauder, *supra* note 16, at 91-92 who describe the process of ‘commensuration’ — a process that involves turning qualities into quantities that share the same metric.

⁴⁹ Carter, Klein and Day, *supra* note 23, at 48.

⁵⁰ ICC OTP, *Strategic Plan | 2016-2018*, 6 July 2015, § 107.

⁵¹ *Ibid.*, at 5, § 1.

affected by a single case, i.e. the contempt case *Bemba et al.* — where each of the five defendants received approximately double the number of total charges that the Office had requested for all other proceedings in the same period.⁵² Hence, about 90% of all charges presented by the Prosecutor in the time period from June 2012 to June 2015 pertain to one single case concerning multiple offences against the administration of justice.

This is not to argue that output indicators are of no value. Indeed, even when further explanation is needed, the OTP's pursuit of positive prosecution outputs is one of its core endeavours. Furthermore, as an international court, the ICC is in a reasonable state of infancy and the Prosecutor's desire to improve prosecution results can be characterized as wanting to put 'legal runs on the board.'⁵³ Successful 'runs' demonstrate the Prosecutor's ability to prosecute defendants through to conviction and, in the process, help improve her procedural conduct at all stages in future cases.

Rather, the more fundamental problem with the OTP's performance indicators is the lack of attention shown to measuring prosecution effectiveness in terms of their specific outcomes (i.e. desired effects on conflict-affected populations that help end the impunity of perpetrators and contribute to the prevention of crimes). This problem can be expressed as the gap between what *ought* to be measured with what *can* (accurately) be measured. In organizations that pursue complex goals, this gap is often articulated with frustration best captured by the expression 'only what is measurable matters.'⁵⁴ The OTP has fallen into this measurement trap, with its failure to measure prosecution effectiveness explained by a combination of outcome uncertainty and the impossibility of unequivocally attributing certain outcomes to its own prosecutorial performance.

However, beyond measurement difficulties, two further concerns explain the lack of attention towards prosecution effectiveness: economy and efficiency.

A. Economy and Inputs

⁵² Decision pursuant to Article 61(7) (a) and (b) of the Rome Statute, *Bemba et al.* (ICC-01/05-01/13), Pre-Trial Chamber II, 11 November 2014. In combination, the five defendants in the case were charged with over 200 offences against the administration of justice under Art.70 ICCSt.

⁵³ P. Clark, 'Law, Politics and Pragmatism: The ICC and Case Selection in the Democratic Republic of Congo and Uganda', in N. Waddell and P. Clark (eds), *Courting Conflict? Justice, Peace and the ICC in Africa* (Royal African Society, 2008) 37-46, at 44.

⁵⁴ See for example G. Bevan and C. Hood, 'What's Measured is What Matters: Target and Gaming in the English Public Health Care System', 84 *Public Administration* (2006) 517-538.

The aim of ‘economy’ is to drive performance towards the reduction of costs (or inputs) or, simply, towards minimising the consumption of resources.⁵⁵ However, in most organizations economy is rarely an end itself, but is usually combined with a commitment to an acceptable benchmark of output quality.⁵⁶ In this regard, this section discusses the OTP’s performance indicators within an economic mind-set. Whilst performance indicators based on economic considerations fulfil (arguably desirable) accountability and audit purposes, they may also produce negative consequences for the Prosecutor’s professional identity. In turn, such consequences further explain the lack of attention with regard to measuring prosecution effectiveness.

Seeking economy while maintaining the quality of outputs is often associated with a ‘hard’ form of managerialism, which holds that reducing the amount of available resources can increase their relative value (i.e. avoiding waste) and, thus, produces better results.⁵⁷ Typically, this requires aggressive monitoring by a management and/or funding body that exerts accountability for those results but simultaneously can place a financial restraint on the organization.

In the ICC’s case, this body would be the Assembly of States Parties (ASP), which elects and can formally remove the Prosecutor.⁵⁸ Critically, the ASP also provides management and budgetary oversight of the Court, including by approving the OTP’s budget.⁵⁹ Arguably, then, there is a clear relationship of accountability between the ASP and OTP. Nonetheless, in the context of financial prudence, the relationship is not as ‘hard’ as in those organizations requiring the achievement of a financial bottom-line (i.e. final net income/profits). Rather, the better description would be that the ASP acts as a principal (or even trustor) with a more indirect influence over a trustee OTP actor.⁶⁰ Whichever the characterization, the effect of the said financial influence has culminated in the ASP’s

⁵⁵ A. Flynn, A. Gray, W. Jenkins and B. Rutherford, ‘Making Indicators Perform’, 8 *Public Money and Management* (1998) 35-41, at 38.

⁵⁶ Carter, Klein and Day, *supra* note 23, at 37.

⁵⁷ Osborne and Gaebler, *supra* note 5, at 15 and 19-20.

⁵⁸ The ASP has the ability to elect or remove the Prosecutor in case of serious misconduct, for a serious breach of his or her duties, or if unable to exercise the functions required by the Statute. See Art. 46(1) ICCSt. See also A.M. Danner, ‘Enhancing the Legitimacy and Accountability of Prosecutorial Discretion at the International Criminal Court’, 97 *American Journal of International Law (AJIL)* (2003) 510-552, at 524.

⁵⁹ Art. 112 ICCSt. (Assembly of States Parties) and Art. 115 ICCSt. (Funds of the Court and of the Assembly of States Parties).

⁶⁰ K.J. Alter, ‘Delegation to international courts and the limits of re-contracting political power’, in D.G. Hawkins, D.A. Lake, D.L. Nielson and M.J. Tierney (eds), *Delegation and Agency in International Organizations* (Cambridge University Press, 2006), 312-338 at 321 and K.J. Alter, ‘Agents or Trustees? International Courts in their Political Context’, 14 *European Journal of International Relations* (2008) 33-63.

demand for a ‘Basic Size’ OTP model.⁶¹ The Basic Size model seeks to provide States Parties with a more predictable basis for budgetary planning, while providing the OTP with a model for managing its financial affairs in view of long-term sustainability. This model depends on financial forecasts of outputs, based on assumptions of their success at various stages from preliminary examination through to trial.⁶²

Moving to the issue of costs, the OTP can be described as the Court’s ‘engine-room’ and is a significant component of the Court’s budget. The OTP’s caseload also determines the financial burdens of other organs, most evidently the Chambers and the Registry. For instance, in 2017, the OTP will be reserved more than €46 million from a proposed Court budget of approximately €147 million.⁶³ As is now well known, states have been notoriously slow in fulfilling financial commitments to the Court, with some remaining in long-term arrears.⁶⁴ For example, between 2009 and 2013 the overall budget growth for the Court was minimal and even shrank in real-terms in 2011.⁶⁵ However, the relative flat lining of the ICC’s budget has been accompanied by the OTP’s accumulating caseload, which from 2009 to 2013 effectively doubled.⁶⁶

In short, the OTP suffers from the aforementioned hard form of managerialism. The ASP exerts the pressure of accountability but in spite of the OTP’s increased workload, it did not proportionally increase OTP funding. This has created tensions in the ASP-OTP relationship. For instance, former Prosecutor Luis Moreno Ocampo publically criticized the UK, France and Germany for having increased the Court’s workload by voting for the referral of the *Libya* situation at the United Nations Security Council, while at the same time advocating to limit the ICC’s budget.⁶⁷ There have also been regular warnings from the

⁶¹ ICC ASP, *Programme Budget for 2015 ...*, ICC-ASP/13/Res.1, 17 December 2014, at § I.3. See also ICC OPT, *Strategic Plan / 2016-2018*, 6 July 2015, §§ 63-70.

⁶² ICC ASP, *Report of the Court on the Basic Size of the Office of the Prosecutor*, ICC-ASP/14/21, 7 August 2015, at §§ 6-17.

⁶³ ICC ASP, *Proposed Programme Budget for 2017 of the International Criminal Court*, ICC-ASP 15/10, 17 August 2016, § 1. The precise figures are €147,250,700 and €46,280.20 respectively.

⁶⁴ See for example, ICC ASP, *Report of the Committee on Budget and Finance on the work of its twenty-seventh session*, ICC-ASP 15/5, 28 October 2016, §§ 16-21. As of 31 October 2016, the current amount of outstanding arrears was in the region of €34 million. See *Report of the Bureau on the arrears of State Parties*, ICC-ASP 15/28, 10 November 2016, § 12. For a discussion of the current shortfall in money see E. Evenson and J. O’Donohue, ‘States should use ICC budget to interfere with its work’, 23 November 2016, *Open Democracy*, available online at <https://www.opendemocracy.net/openglobalrights/elizabeth-evenson-jonathan-o-donohue/states-shouldn-t-use-icc-budget-to-interfere-w> (visited 1 June 2017)

⁶⁵ For discussion on this see S. Ford, ‘How Much Money does the ICC Need?’ in C. Stahn (ed.), *The Law and Practice of the International Criminal Court* (Oxford University Press, 2015) 84-104, at 86.

⁶⁶ *Ibid.*, at 89.

⁶⁷ *Ibid.*

current Prosecutor Fatou Bensouda who, in the past, has protested that the lack of funding will significantly affect the work of the Office and may engender ‘real detrimental results.’⁶⁸

The OTP’s development of performance indicators cannot be divorced from the aforementioned background of restricted funding, combined with an increased workload. As things stand, the ASP’s accountability and funding pressures originate a degree of ‘back-seat driving’⁶⁹ over the OTP, inevitably directing the Office towards demonstrating economy of performance. The ASP’s influence can be readily identified within the OTP prosecutorial strategy. For instance, financial oversight has become increasingly integrated into the OTP’s assessment of its own capabilities and resources.⁷⁰ As a matter of fact, in the latest two OTP strategies, goals are expressly linked to budget assumptions.⁷¹ This may account for the increasing page length of the strategies as they have become progressively more audit, budget and operationally focused, growing from 10 pages in 2006, to 18 pages in 2010, jumping to 43 pages in 2013 and, finally, to the 60 pages of the current strategy.⁷² In addition, as if to demonstrate the pressing importance of economy considerations, the term ‘budget’ is used 57 times in the latter two strategies (combined), after appearing only twice in the first two strategies.

The OTP’s concern about the economy of its activities is not misplaced, but it certainly produces discernible effects on its performance measurement. First, economy risks shaping the purpose of performance measurement into a simple ‘reward-or punish’ exercise. This comes with the belief that hard numbers are the basis for positive external judgment and the search for funding.⁷³ Second, and relatedly, economy pushes performance indicators to produce — as much as possible — hard data, for the purpose of auditing. As argued in the literature, auditing and the development of performance indicators are mutually constitutive.⁷⁴ The development of performance measurement tends to conform to the image

⁶⁸ F. Bensouda, *Address at First Plenary: Fifteenth Session of the Assembly of States Parties*, 16 November 2016, available online at https://asp.icc-cpi.int/iccdocs/asp_docs/ASP15/ASP15-Opening-Statement-Prosecutor-ENG.pdf (visited 31 May 2017), at 8.

⁶⁹ Carter, *supra* note 7, at 138.

⁷⁰ Generally speaking, the 2012 and 2015 prosecution strategies exhibit the features of a ‘SWOT’ analysis. See indicatively, A.C. Hax, ‘Redefining the concept of strategy and the strategy formation process’, 18 *Planning Review* (1990) 34-39, at 37.

⁷¹ See ICC OTP, *Strategic plan June 2012-2015*, 11 October 2013, Annex A and *OTP Strategic Plan | 2016-2018*, 6 July 2015, § 41 and Annex 4.

⁷² The current Strategic Plan includes a 20-page annex discussing the results of the previous Strategic Plan in 2012-2015.

⁷³ E. Ferlie and K.J. Geraghty, ‘Professionals in Public Service Organizations: Implications for Public Sector Reforming’, in E. Ferlie, L. Lynn Jr and C. Pollitt (eds), *The Oxford Handbook of Public Management* (Oxford University Press, 2007) 422-445, at 431.

⁷⁴ Power, *supra* note 10, at 119.

and practice of auditing.⁷⁵ Thus, both accountability and audit explain why the OTP prioritizes indicators based on prosecutorial outputs: because they are readily amenable to auditable standards of performance.⁷⁶

One should then question why audit-rooted performance measurement excludes or dis-incentivizes the measurement of prosecution effectiveness. Generally, one explanation is that these indicators may generate a powerful (and unanticipated) cultural consequence on the organization's personnel. In short, audit indicators may deeply implant audit values within professional identities.⁷⁷ This may lead personnel to change their vision about who they professionally are and what it is they do.⁷⁸ Taking the Prosecutor, her professional identity is not only that of a courtroom advocate concerned with questions of evidence and procedure, but also one akin to a 'cause lawyer': someone whose prosecutions inevitably come with an activist concern and mission to promote social change in conflict-affected societies.⁷⁹ Current OTP indicators may undermine such 'cause lawyer' identity by shifting her concern towards internal-facing considerations, casting her as a private entrepreneur who views prosecutions as a private commodity (rather than a public good), and positive prosecutorial outputs as essential organizational 'profit'. This identity shift also implies a change in the primary audience for the Office's activities, encouraging a focus on the ASP, which holds the OTP accountable and whose demands the Office must strive to meet.⁸⁰ At its worst, quoting Max Weber's language, the Prosecutor becomes 'bureaucratized' and bound by an officialdom comprised of habitual and hierarchical obedience.⁸¹ Therefore, concerns for accountability and audit deflect the Prosecutor's attention away from the desired outcomes within conflict-affected societies.

This article is not suggesting the Prosecutor should not be concerned with economy. That would be impossible and unreasonable, especially in times of financial restraints. Nevertheless, economy alarmingly shapes the purpose of performance indicators and, as a

⁷⁵ *Ibid.*

⁷⁶ *Ibid.*

⁷⁷ Power, *supra* note 10, at 115. See also indicatively, C. Pollitt, 'How do we know how good our Public Services are?' in G. Peters and D. Savoie (eds), *Governance in the Twenty-First Century: Revitalising the Public Service* (McGill-Queen University Press, 2000) 119-155; Hood, *supra* note 22, at 8-10; P. Noordhoek and R. Saner, 'Beyond New Public Management: Answering the Claims of Both Politics and Society', 5 *Public Organization Review: A Global Journal* (2005) 35-53, at 36-44.

⁷⁸ Espeland and Sauder, *supra* note 16, at 86.

⁷⁹ The literature on 'cause lawyering' is considerable, but for a concise discussion see K. McEvoy, 'What Did the Lawyers Do During the War'? Neutrality, Conflict and the Culture of Quietism', 74 *Modern Law Review* (2011) 350-384, at 354.

⁸⁰ The OTP's current strategy makes clear that performance indicators 'also assist the Office to report on its performance to its stakeholders'. See *OTP Strategic Plan 2016-2018*, 6 July 2015, § 103

⁸¹ H.H. Gerth and C.W. Mills (ed.), *From Max Weber: Essays in Sociology* (Routledge, 1995), at 229.

consequence, the Prosecutor's professional identity and the primary audience for her performance can be altered, explaining the neglect in measuring prosecution effectiveness.

B. Efficiency and Outputs

Similar considerations can be derived from an analysis of the impact of the second 'E', i.e. efficiency, on performance indicators.

Efficiency is naturally related to economy. It is concerned with the ratio of inputs to outputs and the desire to gain the maximum possible output from a given input.⁸² The relationship of outputs to inputs is often captured by the notion of 'value for money', where 'value' is still understood as a number reflecting the quantity and quality of outputs.⁸³ This notion tends to resonate in public debate about the Court's performance. For instance, it is hardly surprising that headlines regularly decry the Court's return of convictions from 15 years of operation, particular when its expenses run to well over \$1 billion dollars.⁸⁴ Efficiency is not solely assessed quantitatively by reference to sheer numbers, but also qualitatively: only positive numbers reflect the Office's proficiency.⁸⁵ This section considers how the OTP's indicators privilege an increase in the number and quality of prosecutions, drifting further apart from the measurement of prosecution effectiveness.

It is self-evident that output indicators motivate increasing numbers of outputs. Such indicators inherently exert pressure towards positive (prosecution-leaning) conclusions when the Office faces the dilemma of whether to proceed in a given situation or case. For instance, this happens when, subsequent to an investigation and pursuant to Article 53(2)(a-c) ICC Statute, the Prosecutor must decide: whether there is a sufficient legal and factual basis to seek an arrest warrant or a summons to appear; whether the ensuing case would be admissible; and whether a prosecution would be in the 'interests of justice'. Logically, the extent of available evidence determines the decision to charge a suspect, and thus evidence is the strongest countervailing restraint on simply pursuing an increase in the number of positive charging decisions. On the contrary, the OTP's discretion not to prosecute 'in the interests of justice' is a much weaker restraint. As vague as the expression is, the OTP is on

⁸² Carter, Klein and Day, *supra* note 23, at 38. The dictionary definition of efficiency refers to maximum productivity with minimum wasted effort, or the ratio of useful work performance to the total energy expended. See Stevenson (ed.), *supra* note 2, at 561.

⁸³ Power, *supra* note 10, at 44.

⁸⁴ See for example '13 years, 1 billion dollars, 2 convictions: Is the International Criminal Court Worth it?' DW, 27 January 2016, available online at <http://www.dw.com/en/13-years-1-billion-dollars-2-convictions-is-the-international-criminal-court-worth-it/a-19006069> (visited 1 June 2017).

⁸⁵ Efficiency is so interpreted in Stevenson (ed.), *supra* note 2, at 561.

record stating that there is a general presumption in favour of prosecution, and recourse to this clause would be exceptional and as a last resort.⁸⁶

Perhaps, the strongest factor mitigating the pursuit of increasing numbers of prosecutions is the principle of complementarity of ICC jurisdiction.⁸⁷ In particular, results indicators seem to be incompatible with the concept of ‘positive complementarity’, i.e. an interpretation of the principle according to which the OTP does not only act subsidiarily to national criminal proceedings, but actively promotes them.⁸⁸ In this regard, the OTP has previously declared that the number of cases brought to trial at the ICC should not be the sole or even decisive measure of its effectiveness.⁸⁹ Former Prosecutor Luis Moreno Ocampo went even one step further, stating that the very *absence* of cases would demonstrate the Court’s effectiveness, as it would imply that national authorities are undertaking their own prosecutions.⁹⁰

This confusion could be resolved by resorting to ‘positive complementarity’ performance indicators, measuring the OTP’s impact on national proceedings. However, hitherto, the OTP has failed to develop any. Thus, so long as the OTP’s performance indicators privilege prosecution results, they provide an additional reason for the Prosecutor to interpret complementarity as a competitive ‘grab’ for jurisdiction.⁹¹ In short, the indicators will motivate an aggressive interpretation of complementarity allowing more cases to be considered as admissible, and thus potentially prosecuted.

Turning to quality, while results indicators encourage higher output numbers, they may generate concern as to the final quality of those outputs. The zeal to increase prosecution numbers may affect the Prosecutor’s ability to make an independent and objective assessment

⁸⁶ ICC OTP, *Policy Paper on the Interests of Justice*, September 2007, at 9. See also, albeit in relation to not opening an investigation, Decision on the request of the Union of the Comoros to review the Prosecutor’s decision not to initiate an investigation, *Situation on the Registered Vessels of the Union of the Comoros, the Hellenic Republic and the Kingdom of Cambodia* (ICC-01/13), Pre-Trial Chamber I, 16 July 2015. See also M. Varaki’s article in this Special Issue of the *Journal*.

⁸⁷ Art. 17(1)(a) ICCSt. (Issues of Admissibility), and following.

⁸⁸ There is significant literature on the concept of positive complementarity including its potential incorporation of efforts to build domestic judicial capacity. See, for example, W. Burke White, ‘Implementing a Policy of Positive Complementarity in the Rome System of Justice’, 19 *Criminal Law Forum* (2008) 59-85; J. Tillier, ‘The ICC Prosecutor and Positive Complementarity: Strengthening the Rule of Law’, 13 *International Criminal Law Review* (2013) 507-591; O. Bekou, ‘The ICC and Capacity Building at the National Level’, in C. Stahn (ed.), *The Law and Practice of the International Criminal Court* (Oxford University Press, 2015) 1245-1258.

⁸⁹ ICC OTP, *Report on Prosecutorial Strategy*, 14 September 2006, § 14 and ICC OTP *Prosecutorial Strategy 2009-2012*, 1 February 2010, § 79.

⁹⁰ L.M. Ocampo, ‘Statement made at the ceremony for the solemn undertaking of the Chief Prosecutor’, 16 June 2003, at 3, available online at <http://www.iccnw.org/documents/MorenoOcampo16June03.pdf> (visited 31 May 2017).

⁹¹ On the argument that complementarity acts like a competition for jurisdiction see S.M.H. Nouwen, *Complementarity in the Line of Fire: The Catalysing Effect of the International Criminal Court in Uganda and Sudan* (Cambridge University Press, 2013), at 340.

of facts, evidence and law. After all, it is the quality of these assessments that provide the best prospect for success, increasing the chances of achieving rulings in the Prosecutor's favour.

Can, then, current performance indicators equally bolster the quality of prosecution outputs? One should take into account that these indicators have been developed in an environment that is increasingly recognising the importance of quality over quantity, as plainly acknowledged in the Plan for a Basic Size OTP.⁹² Furthermore, the OTP has explicitly accommodated quality within its current indicator regime. The OTP has recognized that its strategic indicators clearly reflect the quality of investigations and prosecutions, since they attest the success rate of prosecutorial activities.⁹³ Moreover, operational indicator n. 3 measures the 'pattern of judicial findings on how the Office conducts its preliminary examinations, investigations and prosecutions (particularly at the Appeals Stage).'⁹⁴ Hence, the current performance indicators are also focused on the need for the Prosecutor to be meticulous and demonstrate probity because, otherwise, the prospects of successful prosecutions would diminish.⁹⁵

Generally, output-focused performance measurement is embedded in resource-limited environments, where economy and efficiency are primary concerns.⁹⁶ As explained above, the OTP's performance indicators privilege exclusive concern for the number and quality of prosecution outputs.⁹⁷ The adoption of these efficiency-based indicators has further contributed to sliding away from measuring prosecution effectiveness. Nevertheless, the lack of effectiveness measurement is also explained by its very complexity, as illustrated below.

C. Effectiveness and Outcomes

Effectiveness describes the degree to which an organization's outputs translate into an intended outcome. In other words, effectiveness is concerned with the consequences of a given output, including its short, medium and even long-term effects on a target population or

⁹² *Report of the Court on the Basic Size of the Office of the Prosecutor*, ICC-ASP/14/21, 7 August 2015, § 4.

⁹³ ICC OTP, *Strategic Plan / 2016-2018*, 6 July 2015, § 106.

⁹⁴ *Ibid.* One might add that operational indicators nn. 4 and 5 (concerned with compliance with policies and standards and the rather obscure 'quality of interaction with the OTP') also attest to the quality of the OTP's activities.

⁹⁵ The OTP's diligence in case preparation is likely to be stressed in order to avoid negative assessments by the Chambers, that might lead to rulings against the Prosecutor. See for example, Decision on the Confirmation of Charges, *Mbaruishmana* (ICC-001/04-01/10), Pre-Trial Chamber 1, 16 December 2011, §§ 82, 110 where the Chamber was critical of the Prosecutor's specificity in drafting the charges (perhaps due, in part, to the Prosecutor's extensive reliance on evidence in *Human Rights Watch* reports). This contributed to the Chamber refusing to confirm the charges.

⁹⁶ Pollit, *supra* note 1, at 161-163.

⁹⁷ All criminal prosecutors, national or international, are likely to view 'win-loss records' as central to the evaluation of their own performance. On this see C. McCoy, 'Prosecution', in M. Tonry (ed.), *The Oxford Handbook of Crime and Criminal Justice* (Oxford University Press, 2013) 663-696, at 688.

group.⁹⁸ These outcomes are achieved by influencing target audiences — by changing attitudes, behaviour, perceptions and conditions.⁹⁹

In business and public administration contexts, effectiveness is often cast in terms of achieving an outcome of ‘end-user satisfaction’. As such, the OTP can only measure such outcomes after having identified its relevant audiences. Normally, ICC prosecutions speak to several audiences simultaneously, from abstract entities (such as the ‘international community’) to concrete constituencies (e.g. specific institutions, civil society and other state or non-state actors). Yet, the most relevant audiences are those within which the Court’s goals should be empirically achieved. These goals include ending impunity, preventing further criminality and, more in general, promoting stability, peace and reconciliation.¹⁰⁰ Thus, international criminal proceedings are of critical interest for audiences of potential perpetrators, victims and conflict-affected populations.

However, herewith lies the root cause for the OTP’s absence of effectiveness measurement: its lack of clarity in defining the desired prosecution outcomes. Establishing a set of desired outcomes, *a priori*, requires identifying target audiences and the specific outcomes intended for each such audience.

The OTP has previously recognized the importance of outcomes as the starting point for performance measurement by stating, for example, that ‘[t]o evaluate the (cost-) effectiveness of the Office requires first determining what outcome the Office is supposed to produce; prevention of crimes, complementarity achieved, justice (seen to be) done, etc.’¹⁰¹ However, this excerpt is also revealing of the problem. References to outcomes are broad, generic and scattered in differing directions. There has been little reflection about some necessary and crucial distinctions, like the one between prosecution outcomes and the goals of the Court as a whole, or the distinction between outcomes generated from the Office and those from prosecution outputs. For instance, successful prosecutions may achieve specific outcomes (e.g. positive perceptions of Prosecutor decision-making) that may lead to

⁹⁸ Carter, Klein and Day, *supra* note 23, at 35-38. Carter’s definition accords with that of the dictionary, which outlines effectiveness as the degree to which something is successful in producing a desired result. See Stevenson (ed.), *supra* note 2, at 560. Although not relevant for the present purpose, the concept of ‘impact’ is often used interchangeably with ‘outcome’. For a distinction between the two terms see D. Alesani, ‘Results-Based Management’, in Missoni and Alesani (eds), *supra* note 20, 266-294, at 267.

⁹⁹ Carter, Klein and Day, *supra* note 23, at 36.

¹⁰⁰ Literature on the goals of international criminal law and international criminal justice is vast. See for example; L.E. Fletcher, ‘Violence and Social Repair: Rethinking the Contribution of Justice to Reconciliation’, 24 *Human Rights Quarterly* (2002) 573-639; M.R. Damaška, ‘What is the Point of International Criminal Justice’, 83 *Chicago Kent Law Review* (2008) 329-368 and S. Dana, ‘The Limits of Judicial Idealism: Should the International Criminal Court Engage in Consequentialist Aspirations’, 3 *Pennsylvania State Journal of Law and International Affairs* (2014) 30, 31-109.

¹⁰¹ ICC OTP, *Strategic plan June 2012-2015*, 11 October 2013, § 95.

the achievement of the Court's goals, but those prosecution outcomes are not identical to the Court's goals. Alternatively, the Office's wider operations may achieve an outcome of an increase in national proceedings (i.e. positive complementarity) but, evidently, this is not directly related to its prosecution outputs. The OTP would also need to determine whether the desired outcomes are identical for all prosecutions, or whether they should instead be context-specific, taking into account the defendant identity, the nature of the crimes and their socio-political context. Effectiveness cannot be measured without asking these questions. The quoted excerpt — by simply listing some of the outcomes that the Office is 'supposed' to achieve — reveals a casual circumvention of the deep and complex questions that need to be solved before articulating prosecution outcomes.¹⁰²

This lack of clarity precedes and feeds the OTP's concern that it is not possible to unequivocally attribute the production of a certain outcome to the Office's activities (so-called 'performance ownership'). The Office has made clear that it can only measure what it sufficiently controls.¹⁰³ Then, effectiveness measurement has been avoided because of its 'causation complexity': outcomes can be influenced by a multitude of socio-political factors independent from prosecutorial choices.

Even though reasonable (since effectiveness measurement would be of limited value if its causes cannot be sufficiently determined) this approach meets two possible criticisms. First, performance indicators are always open to a degree of contention and scrutiny. Effectiveness indicators would encourage further scrutiny and generate helpful debate about their possible reading. Second, the challenge of measuring effectiveness is logically preceded by the challenge of defining desired outcomes. At present, it seems like the OTP is relying on an abstract 'causation complexity' as an excuse to avoid genuine thinking about desired prosecution outcomes.¹⁰⁴

Moreover, the current absence of effectiveness measurement is compounded by another concern: the OTP's claim that prosecution results indicators *are* indeed a measure of effectiveness. On one hand, one could argue that the OTP intends courtroom wins as an indicator of the Office's effectiveness, if evidential integrity and respect for due process are

¹⁰² On this see also C. Gevers, 'International Criminal Law and Individualism: An African Perspective', in C. Schwobel (ed.), *Critical Approaches to International Criminal Law: An Introduction* (Routledge, 2014) 221-246, at 239-240.

¹⁰³ ICC OTP, *Strategic Plan / 2016-2018*, 6 July 2015, § 104. This understatement is reflected in the ICC's attempts at developing performance indicators for the Court ('the Court needs to be modest and concentrate on a reduced number of measurable criteria that adequately reflects ... performance of the Court without overburdening the exercise with too many criteria and details'). See ICC, *Second Court's report on the development of performance indicators for the International Criminal Court*, 11 November 2016.

¹⁰⁴ Open Society Briefing Paper, *supra* note 12, at 5.

seen as desired outcomes in themselves. However, on the other hand, one could argue that prosecution wins are not sufficient — in themselves — to express the Office’s effectiveness in fulfilling its broader mandate within affected communities, but they are just a starting point towards such end. More worryingly, the OTP’s claim that sheer prosecution results are an indicator of effectiveness may produce a few perverse effects.

First, the OTP’s claim creates a well-known risk in performance measurement: obliterating the distinction between output and outcome.¹⁰⁵ The fusion between the two is common in public administrations, particularly when organizations remain under the pressures of accountability for, and auditing of, outputs.¹⁰⁶ For the OTP, prosecution outputs are expected to lead to prosecution successes, but that success in and of itself is not always equivalent to desired outcomes. For example, an increasingly positive indicator of ‘qualification pass rates’ may well be indicative of a school’s desired outcomes that lead to a body of well-informed and skilled graduates. However, it would not be wholly determinative of it: achieving the said outcomes would depend on the type of the qualification, on the skill level required to gain it, and on a qualitative assessment of the qualification’s recipients.¹⁰⁷ Such confusion between output and outcome can lead the OTP to lose sight of outcomes, further reducing attention towards them. Moreover, the importance of output performance may be overestimated, at the expense of normative discussions about the most appropriate outcomes of prosecutorial decisions.¹⁰⁸

Second, in the absence of meaningful engagement with outcomes, the OTP embodies precisely what Martti Koskenniemi describes as ‘managerialism’. For Koskenniemi, managerialism is an approach whereby decision-makers’ attention is focused on compliance with certain rules, rather than on the reasons for the very existence of such rules.¹⁰⁹ Consequently, achievement of the intended institutional outcomes is simply assumed due to the mere functioning of the institution. In these environments, the only concern is optimising performance towards concrete results – furthering a generic and undefined purpose – through

¹⁰⁵ Power, *supra* note 10, at 115.

¹⁰⁶ For instance, it is alleged that the onerous accountability regime within education in the United Kingdom, has led to schools being disproportionately focused on positive outputs (i.e. examination results). This has generated frequent criticism that schools are being transformed into ‘exam factories’. See for example M. Hutchings, *Exam factories?: The impact of accountability measures on children and young people*, National Union of Teachers, June 2015, available online at https://www.teachers.org.uk/sites/default/files2014/exam-factories_0.pdf (last visited 1st June 2017). See also W. Mansell, *Education by Numbers: The Tyranny of Testing* (Politico’s Publishing, 2007).

¹⁰⁷ For a discussion on indicators and causality, see D. McGrogan, ‘The Problem of Causality in Human Rights Law’, 65 *International and Comparative Law Quarterly* (2016) 615-644.

¹⁰⁸ Geevers, *supra* note 99, at 238-241.

¹⁰⁹ M. Koskenniemi, ‘The Politics of International Law-20 Years Later’, 20 *European Journal of International Law* (2009) 7-19, at 15-16.

the increasing employment of internal expertise, resources and administrative vocabularies.¹¹⁰ Inasmuch as the OTP's performance indicators are an expression of this version of managerialism, they perversely introduce a degree of myopia about the inherent complexity of achieving outcomes. The fact that current indicators give exclusive priority to prosecution outputs exaggerates their very utility. This is what psychologists would label as 'the law of the instrument': 'it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail'.¹¹¹ The Prosecutor may be tempted to think that, in order to achieve the desired outcomes, it is sufficient to use her 'hammer' (i.e. the power to conduct investigations and prosecutions) and 'strike nails' (i.e. actually carry out successful investigations and prosecutions). However, as the OTP itself has recognized when addressing causation complexity, international criminal prosecutions are not the only means towards achieving those outcomes. On the contrary, for example, prosecutions can work within conflict-resolution strategies and alongside other transitional justice mechanisms.¹¹² Thus, the OTP creates a paradox: on one hand it refuses to engage with outcomes because of perceived causation complexity; on the other, it develops indicators that reduce the outcomes' complexity by implicitly assuming they are a natural extension of prosecutorial success.

To summarize, the absence of performance indicators targeting prosecution effectiveness lies, first and foremost, in the OTP's fundamental lack of clarity in defining their outcomes. Furthermore, by juxtaposing effectiveness with its output indicators, the Prosecutor has merged outputs with outcomes, further valorising the former while devaluing the latter. This not only exaggerates the misconception of international criminal prosecutions as a panacea to the world's problems, but it oversimplifies outcomes and naively treats them as a spillover of prosecutorial successes.

3. Conclusion and Recommendations

Performance measurement has long attracted criticism from many quarters. Andrew Natsios once pithily warned against 'obsessive measurement disorder' — a flawed belief that

¹¹⁰ M. Koskenniemi, 'International Law, Constitutionalism, Managerialism and the Ethos of Legal Education', 1 *European Journal of Legal Studies* (2007) 8-24, at 12-13.

¹¹¹ The expression is contained in A.H. Maslow, *The Psychology of Science: A Reconnaissance* (1966; Maurice Bassett Publishing, 2002), at 15.

¹¹² Since the Court only hears a handful of cases, the desired outcomes can only be achieved through a collective effort involving other entities and extra-judicial activities — i.e. education, diplomacy and, more in general, peacebuilding. See ICC OTP, *Strategic Plan /2016-2018*, 6 July 2015, §§ 104-106.

incessant monitoring and measurement will improve policy effectiveness.¹¹³ In contrast, this article has not argued that performance measurement is futile or, even further, that the OTP's efforts at developing indicators ought to be abandoned. It is inescapable that performance indicators continue to be viewed as indispensable. The OTP has even declared that its current indicators are only a pragmatic start-up and, with more experience, it will refine and potentially expand their number.¹¹⁴ This will likely be the case for the development of Court-wide performance indicators, too.¹¹⁵ Hence, the question is not whether to measure performance, but what to measure exactly and how best to do it.

In this light, this article's intent was to stimulate a conversation about the OTP's performance indicators. Current OTP performance indicators are exclusively output-focused, overlooking the need to measure the OTP's effectiveness in achieving certain outcomes. First, the current regime of indicators risks shifting the Prosecutor's professional identity to that of a private entrepreneur, one who is focused on demonstrating financially auditable improvements to the ASP. Second, the indicators privilege concern for prosecution efficiency in terms of their number and quality. Finally, the absence of effectiveness indicators is fundamentally due to the OTP's lack of clarity in defining the desired outcomes of prosecutions, which in turn precedes the challenge of causation, namely attributing ownership over those outcomes to the OTP alone. The Office's claim that prosecution results also measure its effectiveness, moreover, blurs the distinction between outputs and outcomes. In view of the above, the Prosecutor may consider the following principles as a way to improve future performance measurement towards effectiveness.

First, the OTP could realign its strategic focus, making it less about administrative procedures and more about the socio-political environments where prosecutions are meant to have an impact. The current strategic plan has come to resemble an auditing document, dominated by budgetary and operational orientations. However, one of the very functions of strategies is to help ensure that the demands/needs of target audiences are met.¹¹⁶ As also made clear by the UN Secretary-General, prosecution policies related to conflict-affected

¹¹³ A. Natsios, 'The Clash of the Counter-bureaucracy and Development', Center for Global Development Essay, 13 July 2010, available online at https://www.cgdev.org/sites/default/files/1424271_file_Natsios_Counterbureaucracy.pdf (visited 1 June 2017).

¹¹⁴ ICC OTP, *Strategic Plan /2016-2018*, 6 July 2015, § 109.

¹¹⁵ ICC, *Second Court's report on the development of performance indicators for the International Criminal Court*, 11 November 2016, § 3.

¹¹⁶ D. Alesani and I. Bongiovanni, 'Strategic Thinking and Planning', in Missoni and Alesani (eds), *supra* note 20, 239-265, at 246-248.

societies need to take the socio-political context into account.¹¹⁷ Thus, a realigned strategic focus could include an assessment of situation-specific conditions and the precise challenges therein inhibiting the achievement of desired outcomes.

Second, the OTP could commit to a deeper process of qualitative performance evaluation, as opposed to the current preference for quantitative performance auditing. In this sense, the OTP needs to embrace the ‘empirical-turn’ within international criminal justice.¹¹⁸ There is a developing body of literature evaluating the domestic effects of ICC prosecutions within conflict-affected societies. In spite of such research, to date, the OTP has engaged in little self-evaluation about the consequences of its decisions. This could be an annual process, encouraging critical self-reflection about the progress made towards achieving desired outcomes.

Finally, and most importantly, the OTP could make a resolute effort to identify the desired outcomes for its prosecution activities. This effort could be coordinated with Court-wide efforts to develop outcome or impact-oriented indicators.¹¹⁹ These outcomes would need to be far less broad and general than the Court’s goals of ending impunity or the prevention of crimes. Outcomes need to be narrowly defined and context-specific, with a given time-span for their potential achievement. Their measurement could be supported by ongoing empirical fieldwork and be less focused on numerical data. For example, useful indicators would focus on critical issues of legitimacy, measuring perceptions of the Prosecutor’s decisions or beliefs in the Office’s impartiality.¹²⁰ Of course, such indicators would incur measurement difficulties as perceptions — to name only one problem — intertwine with pre-existing ethnic affiliations. Thus, arguably, these indicators could risk producing unreliable data. That being said, risk is not a sufficient reason not to attempt to develop them. Furthermore, the influence of contributing causes on a performance’s outcome is not necessarily fatal for such performance’s measurement. This article’s suggestion is not to develop outcome indicators that are immune to causation complexity, but to critically establish outcome indicators at the core of performance measurement.

¹¹⁷ UNSC, *The Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies: Report of the Secretary General S/2004 616**, 23 August 2004, at 15.

¹¹⁸ G. Shaffer and T. Ginsburg, ‘The Empirical Turn in International Legal Scholarship’, 106 *American Journal of International Law* (2012) 1-46, at 25-30.

¹¹⁹ ICC, Second Report, *supra* note 112, § 4: Currently, the development of Court-wide performance indicators is occurring in parallel and ‘without prejudice’ to organ-specific indicators such as the OTP’s. However there is undoubtedly an overlap between Court-wide intended outcomes and OTP intended outcomes, and thus the development of performance indicators must reinforce this relationship.

¹²⁰ For an example of such empirical fieldwork (though in relation to the ICTY), see J.N. Clark, *International Trials and Reconciliation: Assessing the Impact of the International Criminal Tribunal for the Former Yugoslavia* (Routledge, 2014).

In his work on bureaucracy, Max Weber once described a bureaucrat unable to ‘squirm out of the apparatus in which he is harnessed’, ‘chained to his activity by his entire material and ideal existence’, ‘a single cog in an ever-moving mechanism which prescribes to him an essential fixed route of march’.¹²¹

The Prosecutor can resist becoming such a chained bureaucrat by engaging in the complex challenge of measuring the prosecutions’ effectiveness in achieving essential outcomes within the audiences she ultimately serves. In doing so, she may help recover much-needed faith in the worthiness of her endeavour.

¹²¹ Gerth and Mills (eds), *supra* note 81, at 228.