



OIOS

Office of Internal Oversight Services

**Monitoring, Evaluation and Consulting Division
(MECD)**

**Inspection of the Use of Client Satisfaction Ratings
and Web Metrics as Programme Performance Measures**

**Report No : MECD-2006-006
Date : 13 April 2007
MECD Team: Arild Hauge
Kristian Andersen
Emily Hampton-Manley
Elizabeth Tullett-Court
Janet Closa**

**Report of the Office of Internal Oversight Services on the inspection
of the Use of Client Satisfaction Ratings and Web Metrics
as Programme Performance Measures**

EXECUTIVE SUMMARY

Planning and budget instructions identify “client benefits” as a key dimension of the accomplishments that the United Nations (UN) Secretariat programmes pursue in reference to their General Assembly-approved results frameworks. A diverse range of client satisfaction measurement practices have entered into programme performance planning, monitoring and reporting. The techniques used to determine client satisfaction appear in many cases to be of relatively poor methodological quality, with insufficient attention to client identification, sampling, use of unbalanced rating scales and questionable inferences about findings. Frequently, a narrow range of services and biased respondent selection is employed to substantiate much broader performance claims. Informal techniques, such as compilation of ‘letters of appreciation’ are, in OIOS’ opinion, not appropriate – but are declining in use. Within DGACM and DPI there are some subprogrammes that have sought adherence to sound methodological standards.

Internet traffic statistics – reflecting use of Secretariat website material, not client satisfaction as such – is increasingly used, but technical entry points to website use analysis is continually changing. The most common item of observation at the UN is webpage ‘hits’, a measure that is vulnerable to manipulation. In testimony to their convenience, both client satisfaction measurement and web traffic statistics are utilized to support a higher number of retroactive result claims than results for which they were originally envisaged as the pertinent performance indicator.

Whilst measurement of client satisfaction is not a gauge of the degree to which the UN fulfils its ultimate objectives, it can capture change one step beyond the delivery of outputs. Also, it holds the potential of offering some degree of comparability – over time and across different programmes and types of service. However, the condition of client satisfaction measurement’s validity is the existence and adherence to minimum standards of methodological rigour, for which surveys represent the main instrument. In that regard, current support facilities are inadequate. Although client satisfaction measurement may, with methodological strengthening, yield relevant performance information, monitoring efforts ultimately need to be complemented by programme evaluation – in order for the UN to understand the cause-and-effect relationships that affect observed positive or negative trends - be it in client satisfaction or in other programmatic performance indicators.

Current practices are ultimately constrained by the absence of a corporate accountability framework for what occurs to bureaucratic outputs – and the results beyond. Whether results are achieved or not matters little to resource allocation and individual performance assessment. Programmes have had the option of specifying or adjusting their

data collection methodologies – and thereby, in effect, their performance targets – after budgets have been approved. These factors, which are features of the overall Secretariat planning, budgeting and performance assessment system in general, have had a direct effect on status of practices pertaining to client satisfaction as well as website traffic usage measurement.

OIOS recommends a number of initiatives to improve the level of methodological rigour that is being vested in client satisfaction measurement practice; including formulation of basic standards, the possible establishment of a corporate technical support and guidance facility, an ex ante vetting process, and the establishment of common organization-wide platforms for online surveys and for website traffic monitoring.

CONTENTS

	<i>Paragraph</i>	<i>Page</i>
Abbreviations		
I. Introduction and objectives.....	1-2	
II. Methodology	3	
III. Findings.....	4-73	
3.1 Clients satisfaction in normative planning framework	4-5	
3.2 Client satisfaction and website usage measures are becoming more prevalent	6-8	
3.3 Client satisfaction and website traffic measures are often used as an afterthought	9-10	
3.4 There are no commonly agreed measures for website traffic	11-14	
3.5 Informal methods and ad hoc feedback are inadequate measures of client satisfaction – but becoming less prevalent	15-17	
3.6 Client satisfaction does not necessarily reflect ultimate programme success	18-19	
3.7 Validity of client satisfaction measurement rests upon sound methodology.....	20-26	
3.8 Technical support is wanted; some already available, but not much used	27-28	

3.9 Some shortcomings are shared with other programme performance data	29-30
IV. Recommendations	31-38

Abbreviations

CPC	Committee on Programme and Coordination
DDA	Department of Disarmament Affairs
EA	Expected Accomplishment
ECA	Economic Commission for Africa
ECE	Economic Commission for Europe
ECLAC	Economic Commission for Latin America and the Caribbean
ESCAP	Economic and Social Commission for Asia and the Pacific
ESCWA	Economic and Social Commission for Western Asia
GA	General Assembly
IOA	Indicator or achievement
ICT	Information and communication technology
IMDIS	Integrated Monitoring and Documentation Information System
OHCHR	Office of the United Nations High Commissioner for Human Rights
OIOS	Office of Internal Oversight Services
PAS	Performance appraisal system
PM	Performance measure
PPBME	Rules and Regulation Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods for Evaluation
PPR	Programme Performance Report
RBB	Results-based budgeting
UNCTAD	United Nations Conference on Trade and Development
UNEP	United Nations Environment Programme
UN-Habitat	United Nations Human Settlements Programme
UNHCR	United Nations High Commissioner for Refugees
UNODC	United Nations Office on Drugs and Crime
UNRWA	United Nations Relief and Works Agency for Palestine Refugees in the Near East

.. organizations will not have a long term future if they do not meet the requirements of their customers¹

For many organizations in the public sector the measurement of customer satisfaction will itself be the measure of success.²

I. Introduction and Objectives

1. The current inspection was triggered by the Office of Internal Oversight Services' (OIOS) ongoing concern with quality of the United Nations (UN) Secretariat's systems and practices of programme performance planning, monitoring and evaluation. In the course of preparing the Secretary-General's 2004-2005 Programme Performance Report (PPR)³, the observation was made of a trend towards increasing reliance on performance measures that refer to client satisfaction as well as statistics on website usage. The inspection was conceived to address measurement of client satisfaction and website usage as features of organization-wide performance planning and management practice. The inspection was thus cross-cutting in scope – and does not comprise an in-depth review of practices at individual Secretariat entities. OIOS needs to emphasize that the subject of the current report is not whether clients are actually satisfied with services provided by the Secretariat – but whether the Secretariat has the means to *know*. The specific objectives of the exercise were to:

- Review trends and current status in use of different approaches to the determination of client satisfaction and website usage as performance measures
- Assess the validity and credibility of current techniques
- Recommend possible improvements to current practice

2. Client satisfaction can most generically be defined as *the perception of a client regarding the degree to which a service provider meets or exceeds his or her expectations*⁴. 'Customer focus' is the first principle of the International Standards Organization's (ISO) quality management standards ISO 9000:2000⁵. Website traffic measurement is a somewhat different issue than measurement of client satisfaction⁶. It involves observing activity pertaining to use of website resources and is derived mechanically, i.e. without clients' volunteering their *opinions*. The association stems from website usage being used as evidence of client satisfaction, appropriately or not.

¹ Rochie, G. et al, *Customer Satisfaction Measurement for ISO 9000: 2000*, Elsevier, (2002).

² Hill, N. and Alexander, J., *Handbook of Customer Satisfaction and Loyalty Measurement*, Gower Publishing, Ltd (2000)

³ A/61/64.

⁴ Treasury Board Secretariat of Canada, *Quality Services – Guide II – Measuring Client Satisfaction*, at http://www.tbs-sct.gc.ca/pubs_pol/ojepubs/TB_O/2QG1-2E.asp (1996).

⁵ See <http://www.iso.org/iso/en/iso9000-14000/understand/qmp.html>. See also Bill Self, Greg Roche, 'Customer Satisfaction Measurement for ISO 9000:2000', Elsevier, ISBN 0750655135.

⁶ The exercise was originally announced, further to Memorandum from Under Secretary-General, OIOS, dated 3 May 2006 to all Department Heads, as two separate inspections, respectively on client satisfaction ratings and on web metrics – both in respect of their use as performance measures.

II. Methodology

3. The inspection process involved a combination of desk research, two attitudinal surveys and stakeholder interviews.
 - (i) Desk research comprised an initial tabulation of all Secretariat programmes' references to client satisfaction and web metrics as performance measures in the results frameworks approved by the General Assembly (GA) for the three biennia 2002-2003, 2004-2005 and 2006-2007, as recorded in the Integrated Monitoring & Documentation Information System (IMDIS)⁷. Desk research also comprised review of non-UN literature regarding utilization of client satisfaction and web metrics measures in country-level public sector management.
 - (ii) Two concurrent attitudinal surveys were administered, during the May-August 2006 period, respectively on client satisfaction ratings (in this report referred to as SCS – survey on client satisfaction) and on the use of web metrics (in this report referred to as SWM – survey on web metrics). The SCS was addressed to all 186 subprogramme managers, and the SWM was addressed to 33 departmental website or IT managers. The two surveys yielded respectively 100 and 52 responses⁸, respectively.
 - (iii) Lastly, in-person and phone interviews were conducted with personnel identified by 22 Secretariat departments⁹ as having responsibilities relevant to the current inspection.

III. Findings

3.1 Client satisfaction in normative planning framework

4. At the Secretariat, the immediate association between client satisfaction and performance measurement follows from the 'results-based budgeting' (RBB) system which has been gradually implemented since 2001¹⁰. The underlying purpose of RBB is for planning and decision-making to be driven by future *effects* rather than the mere historical *efforts* of the Secretariat. RBB brings the articulation of results frameworks (frequently referred to as logframes), built on assumed cause-and-effect relationships, as an entry point to strategic planning, resource allocation and reporting. These logframes are part of the budget fascicles that are presented to and finally approved by the GA. They comprise, for all departments' sub-programmes - a set of objectives, expected accomplishments (EA), indicators of achievement (IOA) and performance measures (PM) pertaining to the two-year budgeting periods. Whilst objectives represent an articulation of the basic longer-term rationale for a subprogramme - usually derived from a formal mandate pertaining to a UN programme - EAs reflect the

⁷ See, <http://imdis.un.org>

⁸ Thus yielding a nominal response rate respectively of 53 and 157%. However, we understand that many of our survey questionnaire were forwarded by recipients to colleagues, thus expanding the respondent 'universe' to an unknown quantity – and rendering the calculation of response rates less relevant.

⁹ Further request for nomination of focal points as per Memorandum from Under Secretary-General, OIOS, dated 3 May 2006

¹⁰ Further to GA Resolution 55/231

outcomes to which a subprogramme will contribute within a given biennium. IOAs are the means of verification for those EAs, and PMs are intended to capture the anticipated degree of change (from baseline to target) within a given biennium.

5. Programme performance planning and assessment requirements are encapsulated by the Rules and Regulation Governing Programme Planning, the Programme Aspects of the Budget, the Monitoring of Implementation and the Methods for Evaluation (PPBME)¹¹ and the instructions that are periodically issued in support of planning and budgeting¹² and performance reporting¹³. The PPBME actually defines “Expected Accomplishments” as centred on client benefits: “*Expected accomplishments...shall identify those benefits or changes expected to accrue to users or beneficiaries*¹⁴ ...”.

3.2 *Client satisfaction and website usage measures are becoming more prevalent*

6. It is evident that reliance on client satisfaction and/or website usage measurement has significantly increased in both nominal and relative terms as a Secretariat practice, with a clear majority of departments now utilizing such measures (within one or more of their subprogrammes). For the 2002-2003 biennium, OIOS found that out of 649 IOAs that were listed in departmental budget fascicles, 73 (11%) made reference to client satisfaction¹⁵ or to website usage (see table 1 below). These were spread across 10¹⁶ out of 32 (31%) programmes reviewed. For the 2004-2005 biennium, the number of references had increased to 115 out of 974 IOAs; relative to total number of IOAs a slight increase (12%), but involving a higher share of programmes, 22 out of 32 (69%)¹⁷.

7. For the 2006-2007 biennium, mid-term records¹⁸ indicate that, across the Secretariat as a whole, the number of such references increased to 201 of the then 992 IOAs (20%), deriving from 28¹⁹ out of the 33 Secretariat programmes (85%)²⁰. The observed trend towards increased reliance on client satisfaction as a performance measure appears likely to continue, as evidenced by 89% of respondents to the SCS expressing the opinion that client satisfaction measurement will, in the *future*, be either ‘important’ (37%) or ‘very important’ (52%).

¹¹ ST/SGB/2000/8, Rule105.4 (a) (iii)

¹² See <http://ppbd.un.org/rbb/>

¹³ See <http://imdis.un.org/>

¹⁴ PPBME Rule105.4 (a) (iii). We note that RBB guidelines for 2008-2009 biennium (<http://ppbd.un.org/bi08/Rbbguide.pdf>) introduces as further refinement, ref p. 26: “The formulation of the result should answer the question “What benefit will accrue to the end-user at the end of the biennium?”.

¹⁵ Based on a word-search followed by elimination of instances of double-counting

¹⁶ DDA, DESA, NEPAD, ECLAC, ESCWA, HCHR, DPI, DM-PPBA, DM-OHRM, OIOS

¹⁷ Additions were: DGACM, UNCTAD, UNEP, HABITAT, ODC, OCHA, DM-OCSS, DM-UNOG, DM-UNOV and DM-UNON

¹⁸ Programmes have had the option of specifying (or adding to) their list of performance measures *after* the GA approved their budget. Although the total number of such measures can thus increase over time, until the end of the biennium, we have considered IOAs and PMs as being most essentially part of the ex ante planning process.

¹⁹ Further additions being: DPKO, OLA, UNHCR, ITC, ESCAP, ECE

²⁰ Including subprogrammes for which separate budget fascicles are issued.

Table 1. Secretariat use of client satisfaction or web use measures in performance plans a reporting

	2002-2003	2004-2005	2006-2007
Number of Indicators of Achievement (IoA)	649	974	992
Total No of References to Client Satisfaction or Web Use	73	115	201
Share (%) of IoAs	11	12	20
Total No of Departments citing Client Satisfaction or Web Use	10	22	28
Share (%) of Departments citing Client Satisfaction or Web Use	(10/32) 31%	(22/32) 69%	(28/33) 85%
Total # of Results Achieved citing Client Satisfaction or Web Use	74	184	n/a
% of Results Achieved citing Client Satisfaction or Web Use	16%	29%	n/a

8. Among techniques, the use of client surveys is predominant – accounting for 82% of the references made to client satisfaction or website usage in the IOAs for 2006-2007, as per table 2 below. Amongst the surveys conducted by SCS respondents, 42% were paper-based and administered in person, 38% were distributed as attachment to e-mails, and only 30% were administered as web-based surveys. OIOS noted that there is no current organization-wide software system for administering online surveys. Several programmes have independently, and in parallel gone through a vendor selection process in respect of procuring software for web based surveys²¹. In other cases, online survey instruments have been designed from ‘scratch’ – based on internal expertise and capacities.

Table 2. Types of measurement techniques in use at the Secretariat – by IMDIS word reference

	2004-2005		2006-2007		Change	
	No.	(%)	No.	(%)	No.	(%)
Survey-based client satisfaction ratings	83	72%	165	82%	+82	+10%
Indicators of Website traffic	21	18%	24	12%	+3	-6%
Informal reviews, Letters of Appreciation, Citations in publications etc	11	10%	12	6%	+1	-4%
Total	115	100%	201	100%	86	+75%

3.3 Client satisfaction and website traffic measures are often used as an afterthought

9. Changing the unit of observation from *ex-ante* IOAs to *ex-post* results statements reveals another dimension to the picture of increasing reliance on client satisfaction and website usage measurement. Whilst the share of *ex ante* IOAs citing client satisfaction or website use was relatively stable in the 11-12% region between 2002-2003 and 2004-2005, the proportion of end-of-biennium *ex post* results statements that referred to such methodologies increased from 16 per cent to 29 per cent. This suggests that programmes, when retroactively making result claims, end up being more dependent on client satisfaction measurement than they envisaged at the beginning of the planning cycle. This was especially pronounced at end of 2004-2005, for which

²¹ ‘Websurveyor’, ‘Survey Monkey’, ‘GMI’, ‘Snapsurveys’ and ‘Questback’ being among the software providers cited by departments

the share of *ex post* results statements referring to client satisfaction or website usage was more than two-and-a-half times the share of *ex ante* IOAs that made such reference. In comparison, for the 2002-2003 biennium, the share of *ex post* statements referring to client satisfaction was only somewhat higher than *ex ante* IOAs²².

10. Likewise, when it comes to web traffic alone, at the *end* of the 2004-2005 biennium, a total of 74 *ex post* results statements referred to web usage – being more than three-and-a-half times the number of *ex ante* IOAs for the same period. This means, again, that web traffic measures are employed in support of many more results than those for which web traffic was originally identified as the pertinent performance indicator. In turn, this suggests that website traffic indicators are found, in hindsight, to be more relevant – or simply more convenient – than envisaged at the beginning of the biennium.

3.4 *There are no commonly agreed measures for website traffic*

11. Web traffic can be measured in many ways, including analysis of number of downloads, hits, unique visitors, page views and other user tracking data. Each approach also comes with limitations in the inferences that can be made from the quantitative findings they yield. Log file analysis can extract accurate ‘*page view*’ or ‘*page request*’ details, breaking down visits to individual website sub-components (i.e. the different pages within a given website), the duration of ‘*visits*’ and the geographical distribution of those who have entered - but may provide an incomplete picture of use, e.g. due to *caching*²³. The measurement of downloads can also be technically challenging. For instance, one needs to know the number of successful, not just requested downloads.

12. Amongst Secretariat programmes, the most common approaches to gauging website traffic is to count hits, followed by downloads, duration of time spent on a web page or study of log files of user patterns²⁴. The use of ‘pop-up’ surveys to obtain more facts about website users and their perceptions about materials perused – i.e. determination of satisfaction - is limited²⁵. Several subprogrammes still make retroactive performance claims (for 2004-2005) and *ex ante* performance plans (for 2005-2006) based on increased volume of ‘hits’. However, *hits* have largely been discredited as a measure of web site traffic²⁶. Because a single web page can contain dozens or more different elements that are separately counted, the use of hits make comparisons meaningless. Also, performance against targets can, in effect, be manipulated by changing

²² Because these observations emanate from statements provided at end-biennium, comparable numbers do not yet exist for 2006-2007.

²³ Entailing that when a web page is being viewed the user may not actually be visiting the website maintained by the content ‘owner’, but a ‘cached’ version stored elsewhere (such as databases of a *search engine*) - and the content owner never acquiring any record of it having been viewed. Conversely, visits made by search engines and ‘*robots*’ looking for content that is not necessarily leading to material being viewed, can also *inflate* the number of entries in log files – thus not showing visitors who actually made use of content. Factoring out the activity of search engines is thus necessary in order to obtain a realistic and accurate picture of web traffic. Likewise, programme-internal access – i.e. staff who access a programme’s own website, e.g. for commonly used documents and materials – need to be factored out in order to arrive at any measure of client use.

²⁴ Respectively in use by 36%, 34%, 9% and 10% of respondent departments.

²⁵ 15% of respondent departments indicate such use.

²⁶ For a more detailed and supplier-independent review of different techniques, see <http://www.computerworld.com/managementtopics/ebusiness/story/0,10801,71989,00.html>

content (e.g. adding photos) – without actually receiving any additional visitors. There are numerous different commercial software products to support website usage or traffic monitoring currently in use by Secretariat programmes²⁷. Use of commercial products is frequently supplemented by programming efforts of the individual subprogrammes.

13. In any case, all automated web traffic monitoring systems yield nominal and largely quantitative data on volume – without giving information allowing inferences about satisfaction of users. Ultimately, the satisfaction of website users can only be determined through supplementary, qualitative techniques, e.g. through interviews, focus groups or more in-depth surveys of website users.

14. The use of internet has become a mainstay of the Secretariat's operations. At the moment, the determination of the web traffic use is fragmented among observation of multiple different technical web parameters. OIOS is unable to prescribe the exact parameters that are most efficacious to future monitoring. The industry 'benchmarks' for what is useful – and possible – to track are continuously evolving. In this respect, it is apparent to OIOS that a degree of flexibility will be needed – to avoid programmes becoming 'locked into' performance targets and measurement techniques determined at the planning stage – but that are either irrelevant or cost ineffective by the time of actual programme implementation.

3.5 Informal methods and ad hoc feedback are inadequate measures of client satisfaction – but becoming less prevalent

15. The use of 'letters of appreciation' as a methodology for measurement of client satisfaction by the Secretariat programmes has declined between 2002 and 2007²⁸. By 2006-2007, letters of appreciation were not included among the planned performance measures, as stated at beginning of biennium. OIOS' assessment is that letters of appreciation are especially vulnerable to subjective analysis. A review of samples submitted to OIOS indicates that letters of appreciation and complaints are very variable in specificity and that there is little standardization – even within individual departments – of the processing, recording and response to such letters. Several programmes wrote that they did not have "specific" or "formal procedures". While they may be used to express sincere gratitude for participation in an event that, in the sender's view, was well-executed – the letters reviewed by OIOS are not clearly relevant to the EAs that Secretariat departments have committed themselves to. As highlighted by one SCS respondent: "such letters measure political appreciation, (and are) not necessarily merit-based, reflecting reality." Also, they "do not measure dissatisfaction". At the end of the day, the writing and recording of letters of appreciation needs to be seen as part of the customs of diplomatic protocol and courtesy – being nice to receive, but nominal in substantive focus and unreliable as a performance indicator.

16. Most programmes (56%) keep record of the number of times they are cited in media, academic journals – and in official records of UN proceedings – and several relate such

²⁷ We note e.g. DDA's use of 'Sawmill', ECLAC's use of 'Web Trans', ESCAP's use of 'Urchin', HABITAT's use of 'Deep Matrix'. Other applications mentioned include 'Webalizer', 'Analogue', 'NetIQ', 'SADE'

²⁸ During the 2002-2003 biennium, letters of appreciation were used in support of performance reporting by DDA, DESA and ECLAC.

information to client satisfaction. Of course, the frequency or volume of such citation does not necessarily indicate client satisfaction – the context of UN mention may, on the contrary, be entirely critical. Similarly, there are cases²⁹ in which *number of participants* in meetings is interpreted as an expression of satisfaction. In this case too, nominal volume of participation cannot necessarily be interpreted as satisfaction with services provided by the UN.

17. OIOS’ assessment is that ‘letters of appreciation’ and other informal sources of feedback are not, in general, adequate as measures of client satisfaction. When expressed by specific clients in respect of specific services provided by specific Secretariat staff – they may be relevant to assessing the performance of those specific Secretariat staff. However, their usefulness as a measure of performance at the overall programme level is, in general, limited.

3.6 Client satisfaction does not necessarily reflect ultimate programme success

18. OIOS recognizes from the outset that, to the extent that client satisfaction relates to attitudes, perceptions and other “proxies” for “real-world” phenomena that the Secretariat seeks to effect, it is a less-than perfect measure of performance. Indication of client satisfaction ultimately represents a measure of performance that should be complemented by triangulation with other types of observations. An example of performance measures that are *more* relevant than client satisfaction – as a single measure of overall programme success – are those used by the UN Relief and Works Agency for Palestine Refugees in the Near East (UNRWA), which is a provider of key social services directly to over 4 million refugees. Although the clients are clearly identifiable and their levels of satisfaction can be gauged, there are actually more objective and substantive measures of performance available. UNRWA is able to report³⁰ on the actual health and educational status of those served by the agency (e.g. school pass rates; infant and maternal mortality; sewerage connection and access to safe water). When data is simultaneously available for a control group (i.e. Palestine refugees *not* served by UNRWA), the differential will be seen as associated with UNRWA – and founded on an evidential basis that is more substantively relevant as a performance measure than perceptions of satisfaction.

19. Nevertheless, OIOS’ assessment is that the notion of client satisfaction does have validity. Secretariat entities vary greatly in the nature of their operations as well as what may be considered appropriate as a measure of their programmatic performance. Many of the UN’s functions are process-oriented, involving global forum-convening and norm-setting rather than direct delivery of services to the public. Objective and uniform indicators of impact, efficiency or effectiveness can be elusive. The ‘real-world’ effects of what the UN does may only materialize over a very long timeframe – and then be difficult to separate from the contribution of other actors and factors. As such, client satisfaction does represent a notion of performance that goes at least one step beyond the measurement of internal bureaucratic activity. All departments provide a service of some kind for which a set of clients can be identified, be it an internal or an external constituency. Client satisfaction ratings also hold the potential of allowing *some* degree of

²⁹ E.g. OLA, Law of the Seas subprogramme

³⁰ E.g. in context of 2004-2005 Programme performance of the United Nations for the biennium 2004-2005, A/61/64, pp. 211-216

comparability across location, types of operation and time³¹. Surveys, in particular, have the advantage of being adaptable to a number of varying environments and can be administered by placement on websites, through email or paper media or through phone interviews.

3.7 *Validity of client satisfaction measurement rests upon sound methodology*

20. Whilst potentially relevant, there are conditions attached to the utility of client satisfaction measurement. Above all, there is a need for a higher degree of consistency and rigour to the methodological foundation of client satisfaction measurement practices. In that regard, an initial set of conditions relates to definitional clarity in terms of: a) alignment between queried services and expected programme accomplishments, b) existence of an identifiable and legitimate client constituency, and c) appropriate techniques for determining satisfaction. These concerns, in turn, bring focus to the imperative of minimizing survey error. Firstly, there may be ‘non-response’ bias, i.e. those who provide feedback being representative only of those who have received service – not those who are *meant to* receive service³². Secondly, those who do *respond* may not be typical of those who have received service – only those who have strong positive or negative feelings. Thirdly, those who do respond may not be truthful – and instead provide answers that they think are wanted or that they think they themselves will benefit from. Fourthly, there may be measurement errors – whereby inaccuracies follow from the way questions are framed or responses tabulated³³. OIOS notes, albeit without having conducted an in-depth review of individual instruments, that DGACM and DPI are among the few departments that have sought to maintain efficacious survey methodology.

21. OIOS found a number of instances where there is a mismatch between the EAs that are being pursued and the actual services about which expressions of satisfaction from clients have been sought or expressed. An example would be an EA framed as “effective implementation of outcomes of (global conference)³⁴” being validated, at end-biennium, by reference to participants’ satisfaction with support provided to meetings of a particular commission or committee backstopped by the subprogramme in question. A similar example would be that an EA on ‘enhanced policy dialogue on trade practices and regulatory framework’³⁵ is evidenced by satisfaction expressed in respect of a particular forum meeting that has been held. In these cases, the mismatch is most importantly one of magnitude – i.e. that the service about which satisfaction is expressed is little more than a narrow ‘slice’ of the EA – and thus not sufficient as evidence of progress towards the much ‘bigger’ EA.

22. Satisfaction is itself a complex issue – and may comprise perceptions about the degree to which a service is pertinent to a respondent’s needs; their feelings about whether service delivery

³¹ As an example of a client satisfaction measurement methodology that has been applied to multiple contexts, covering both private and public services, see the *American Customer Satisfaction Index*, at <http://www.theacsi.org/overview.htm>.

³² See e.g. Everett, S., 2000 respondent satisfaction measurement. Council for Marketing and Opinion Research (CMOR). Port Jefferson, NY (2000) www.cmor.org or Fletcher, J. and Schmidt, D., *Measuring response bias in survey research*. A paper presented at AAPOR, annual conference, May 16-19, 2001

³³ There are numerous standards and definitions of survey error, see e.g. US Census Bureau, <http://www.census.gov/>, or OECD, <http://stats.oecd.org/glossary/index.htm>.

³⁴ Example from DESA, Sustainable Development subprogramme, 2004-2005

³⁵ Example from ECE, Trade and Development subprogramme, 2006-2007

has been well executed – and may or may not be expressed relative to a particular set of expectations. At the Secretariat, when clients satisfaction perceptions are sought, it is general satisfaction that is most frequently (81%) queried, although timeliness, quality, technical expertise are also raised by a majority of SCS respondents as being focused upon. Clarity in these regards is indeed instrumental to validity and utility of client satisfaction measurement as a tool to instigate improvements in service delivery.

23. It is not clear to OIOS that those managers who conduct client satisfaction surveys, in general, make appropriate efforts to minimize sampling bias. For some services, e.g. public documents, there is no finite universe of clients – and response rates to surveys are often unknown. Whilst the direct recipients of services from the UN may be possible to identify, *their* satisfaction is not necessarily the same as that of the true ‘population’ or ‘universe’ of clients most appropriately associated with the EA. For instance, the people who attend UN conferences or meetings, perhaps with funding provided by the UN itself, may be entirely satisfied with their participation – without their attendance ever translating into any benefits to the intended, ultimate beneficiaries of the UN’s work – whose satisfaction would more closely mirror the stated EAs.

24. From OIOS interviews there is some indication that client satisfaction surveys are administered by selecting a service (or event) that is generally considered a success – and thus not be typical of the full range of services needed to make progress towards an expected accomplishment. Along the same lines, there are several cases of satisfaction ratings being expressed in reference to unbalanced scales³⁶ – whereby number and labelling of response options are tilted towards yielding favourable ratings³⁷. OIOS notes that, in administering surveys, practices for maintaining respondent anonymity also vary. Lastly, we have found no subprogrammes that have made any detailed description of their client satisfaction *methodology* publicly available. All survey and public opinion researchers have an obligation to provide certain minimal information about how research was conducted, in order to allow consumers of survey results an adequate basis for judging the reliability and validity of results reported. At the UN Secretariat, such a practice would be in line with the Secretary-General’s reform proposal³⁸ and suggestions for improving public access to Secretariat information³⁹.

25. An important risk to the utility of questionnaires as method for obtaining client satisfaction data, raised by many inspection interlocutors, is the possibility of ‘survey fatigue’ – i.e. that clients (whether internal or external) will avoid responding to questionnaires – not because the questionnaires are poorly designed or because they have no opinions to offer, but because they get so many of them⁴⁰. Whilst agreeing that this is a risk, OIOS notes that it is also, in part, a question of survey design and management – i.e. that too many departments ask too many and too general questions from a client group that has not been too well defined. We

³⁶ e.g. where more calibration is presented for positive than negative perceptions – for instance by offering a range of response options comprising ‘excellent’, ‘good’, ‘satisfactory’ and ‘poor’.

³⁷ See e.g. Sangster, R. and Willits, F., *Evaluating Numeric Rating Scales: Replicated Results*, United States Department of Labor Bureau of Labor Statistics, 2001, at <http://www.bls.gov/ore/pdf/st010120.pdf>

³⁸ See proposal 19 of A/60/692 and Corr.1, ‘Investing in the United Nations for a Stronger Organization Worldwide’

³⁹ A/60/846/Add.4

⁴⁰ For transparency we need to make public a complaint made by some inspection interlocutors, namely of being recipients of three different simultaneous surveys from OIOS’ MECD division alone.

believe this to represent a factor that strongly lends support to the need for organization-wide standards and coordination.

26. OIOS's general consideration is that there is not sufficient rigour and discipline to the measurement of client satisfaction and website usage. OIOS notes that there has, in general, been little progress towards the earlier call, by the CPC⁴¹, for 'consistent standards' for survey conduct. Finally, OIOS notes that, although client satisfaction measurement may, with methodological strengthening, yield relevant performance information, monitoring efforts ultimately need to be complemented by programme evaluation practice that addresses the full range of cause-and-effect relationships that affect observed positive or negative trends⁴² - be it in client satisfaction or in other programmatic performance indicators.

3.8 Technical support is wanted; some already available, but not much used

27. Whilst some departments, notably DGACM and DPI, have developed an in-house expertise for the review of client satisfaction and website usage, a large majority of informants, both interviewees and respondents to the SCS the SWM agree that lack of financial and human resources were the greatest obstacles to adequate practices⁴³. Amongst respondents to the SWM, an overwhelming majority also underlined the importance of policies and guidelines. Only one programme⁴⁴ reported receiving any training from ITSD on analysis of web metrics data. We understand Webtrends - the software in use by ITSD - to be a *de facto* industry standard, but note that its use is limited. One critical factor to that effect is that use of ITSD services involves a charge or cost for which other Secretariat programmes have no budget allocation.

28. OIOS has reviewed the body of technical guidance available at DM's RBB website⁴⁵, and found the materials to be relevant - but not to have been used much⁴⁶. Not a single inspection interlocutor referred to the guidance materials available. OIOS also notes that little, if any, revision has been made since 2003. Whilst materials highlight pertinent principles - little reference is made to cases of potential individual good practices⁴⁷ that already exist. Subprogramme staff wish to have more direct, practical 'show-me-how-to' type of advice. There are currently no staff in DM or elsewhere in-house available to provide the hands-on assistance that subprogramme managers express a need for. Overall, 62% of respondents to the SCS, agreed that there is a need for common guidelines and minimum standards for conducting surveys.

⁴¹ As per A/59/16 (Supp), para. 27

⁴² See e.g. OIOS' report A/61/83 'Strengthening the role of evaluation and the application of evaluation findings on programme design, delivery and policy directives' and OIOS annual report (A/60/346 and Corr.1)

⁴³ Respectively by 54 and 55% indicating that factor as a 'big' or 'moderate' obstacle.

⁴⁴ Department of Disarmament Affairs (Geneva branch)

⁴⁵ See 'Guide to RBB', dated 23 October 1998, <http://ppbd.un.org/bi08/Rbbguide.pdf>, Guide to 'Preparation for Data collection and measurement', <http://ppbd.un.org/rbb/prep.doc>, Guide to 'Collecting Data', <http://ppbd.un.org/rbb/colldata.doc> and Guide to 'Post data collection', <http://ppbd.un.org/rbb/postdata.doc>

⁴⁶ As of 22 January 2007, a total of 7921 visitors was cited on the website itself.

⁴⁷ OIOS found that e.g. DPI conducts surveys with attention being paid to the client identification, sampling, phrasing of questions, and utilization of Likert scales. Efforts to maintain historical data has also been started with an Access database. The efforts of DPI have been self-directed but their experiences could be shared with other departments or subprogrammes.

3.9 *Some shortcomings are shared with other programme performance data*

29. The measurement of client satisfaction, respectively website usage, is subject to the same systemic constraints that apply to the broader enterprise of programme performance management at the Secretariat in general. Perhaps most crucially, programme performance measures, in general, do not have a direct decision-making purpose or accountability implications. This is, however, a characteristic of the RBB budgeting process that is centred upon results that are aimed at, not results that have actually been achieved. When actual performance against budget objectives is reported, through the PPR, budget decisions for the next budget period have already been made. The primary recipient of the PPR, the CPC, does not, in any case, have authority over budget resources. At the same time, the PPBME explicitly state that no information shall be transmitted between the programme evaluation and the personal performance appraisal systems⁴⁸. Likewise, there are no client satisfaction measures that have been included in the indicators to be reviewed by the Management Performance Board⁴⁹. OIOS' findings reaffirm the general picture noted by the Secretary-General⁵⁰, that "*The existing systems for reporting and evaluating the performance of programmes have no practical impact on future plans and resource allocation decisions*".

30. Similarly, on the methodology front, specification of IoAs and PMs has not necessarily been disciplined by realism in availability of underlying data. The biennial budget process has not required that data collection methodologies for IoAs and PMs be specified by the time of budget approval⁵¹. In many cases, performance indicator methodologies specified in IMDIS are aspirational – being things that *could* or *should* be done to determine performance – without it being clear whether they *will* be used or not. Programme managers can, in effect, modify their own performance targets during the biennium. Managers will always wish to put their own performance in the best possible light – and therefore choose measures accordingly. On the other hand, we note that there are a number of instances where client satisfaction has been recorded as 100% - as both baseline and target – thus leaving the measure of little utility to improving performance.

IV. Recommendations

31. The status of the UN Secretariat's use of client satisfaction as a measure of performance cannot ultimately be seen in isolation from the underlying systems and practices of programme performance planning, budgeting and reporting. In that regard, OIOS' analysis and recommendations should be seen in the context of the Joint Inspection Unit's (JIU) observations on results-based management at the Secretariat⁵² as well as OIOS's own assessment of the need

⁴⁸ PPBME rule 107.3 (e)

⁴⁹ As per A/61/319, 'Management Performance Board'

⁵⁰ As per A/57/387, 'Strengthening of the United Nations: an agenda for further change', para 164

⁵¹ As of 1 February 2004, i.e. after beginning of 2004-2005 biennium, data collection plans had been specified for only 25 per cent of the Secretariat's total range of performance indicators. By 31 January 2005, that share had risen to only 46 per cent. In respect of the 2006-2007 biennium, as of 17 July 2006, indicator methodologies had been specified for 55% of the Secretariat's 1021 IoAs. By 18 January 2007, such methodologies had been specified for 67% of IoAs.

⁵² See JIU/REP/2004/5, 'Overview of the series of reports on managing for results in the United Nations system' and JIU/REP/2006/6, 'Results-based management in the UN in the context of the reform process'.

for strengthening programme performance monitoring and evaluation⁵³. The ongoing reviews of RBM and of the experiences gained with planning and budgeting system requested by the General Assembly⁵⁴ presents an opportunity for placing the current findings and recommendations within a comprehensive set of considerations pertaining to the broader decision-making process at the UN.

All the following recommendations are addressed to the Department of Management.

Recommendation 1

32. ***OIOS recommends that a set of minimum methodological standards for survey conduct be established.*** This may comprise of a definition (and subsequent circulation of guidelines pertaining to e.g.) survey constituency, sampling techniques, presentation of findings and public availability of methodological description – and should be integrated within the format for IMDIS as a description of such methodologies. One item of methodology that can potentially be addressed separately from broader issue of standards and guidance is promulgation of a uniform scale of satisfaction ratings for perception surveys. OIOS believes that practices for use of ‘Likert’ scales⁵⁵ can and should be standardized into a simple and balanced format, e.g. with a five-point scale – for use in all surveys that address ‘strength of attitudes’. (SP-06-006-001) *.

Recommendation 2

33. ***OIOS recommends that ‘letters of appreciation’, informal document review and feedback be discontinued as performance measures.*** Advice to this effect would need to be integrated with instructions for articulation of strategic frameworks, budget proposals and performance monitoring – as well as being integrated into the body of guidance available on an ongoing basis (i.e. websites, manuals etc.) (SP-06-006-002).

Recommendation 3

34. In order to enable implementation of minimum methodological standards as above, ***OIOS recommends that consideration be given to the establishment of an advisory facility for client satisfaction measurement and survey conduct.*** This recommendation may most appropriately be addressed in the context of reviewing broader technical guidance in support of the UN’s overall RBB/RBM system for the planning, budgeting and reporting of programme performance. (SP-06-006-003).

⁵³ E.g. as per pp. 11-18 of A/60/73 and pp. 95 of A/61/64, ‘Programme Performance Report of the UN’.

⁵⁴ As per resolutions 58/269 and 61/245

⁵⁵ Likert scales are frequently (knowingly or not) used for asking a person to select a category label from a list that expresses intensity of attitudes or indicating the extent of disagreement or agreement with a statement.

* An internal code used by OIOS for recording recommendations.

Recommendation 4

35. In order to strengthen the demand for application of good practices as recommended above, ***OIOS recommends that consideration be given to the establishment of a mechanism for vetting, prior to review of individual departmental budget submissions, of the client satisfaction measurement methodologies that are being proposed as a basis for programme performance assessment.*** This would, in turn, involve: a) formulation of criteria for review; b) assigning responsibility for review of methodology; and c) communicating the requirements and approach to the budget applicant departments. (SP-06-006-004).

Recommendation 5

36. ***OIOS recommends that consideration be given to procurement and installation of a common software platform for conduct of online surveys***⁵⁶. This would eventually support convergence in practices, allow for accumulation of organization-wide statistics on client satisfaction, and could lead to substantial scale economies in vendor selection and maintenance. A common online survey platform would, however, not merely be a software consideration – it would also need to be seen in context of broader advisory capacity and support facilities. (SP-06-006-005).

Recommendation 6

37. ***OIOS recommends the establishment of an interdepartmental advisory group or task force on website traffic monitoring be initiated.*** This body should be tasked, above all, with articulating and periodically updating a body of good practice pertaining to the technical parameters of website traffic monitoring. It should also review the functional conditions and costs of establishing a common, organization-wide software platform for such website traffic monitoring. (SP-06-006-006).

38. Finally, OIOS notes that, further to its communication⁵⁷ of early findings from the current inspection exercise, DM has included a number of pertinent revisions to its budget instructions for the 2008/2009 biennium⁵⁸. However, the current final inspection report brings further specificity to findings and recommendations – and to the actions that need to be taken by the Secretariat. Report of the Committee on Relations with the Host Country (resolution 61/41).

⁵⁶ We recognize that whilst ITSD may need to have technical responsibility for such installations, its functional parameters would need to be determined by user programmes.

⁵⁷ As per e-mail exchange 7 September 2006.

⁵⁸ Ref Para 22 page 11-12 of 'Instructions: Proposed Programme Budget for the Biennium 2008-2009', intranet: <http://ppbd.un.org/bi08/>.

