

UPU-agreed measurement systems External audit 2022

Universal Postal Union
International Bureau

Audit report
February 2023



Table of contents

1. Executive summary	2
2. Scope of our work	6
3. Audit methodology and process	7
4. Audit results.....	9
A1 Rating criteria	20

1. Executive summary

1.1. Scope of work

The UPU GMS has been running Quality of Service measurements since 2009, starting with 21 designated operators (DOs). This number rose over the years reaching 60 DOs in 2022 that participated in the UPU Quality of Service link to terminal dues (commonly referred to as UPU QS link). Similarly, the International Post Corporation (IPC) has been running the UNEX for the UPU terminal dues (UNEX UPU TD) measurement system, for which 15 DOs were measured and participated in the QS link in 2022. Having two UPU-agreed measurement service providers (MSPs) measuring 75 designated operators for the purposes of the QS link calls for transparency and reliability in the measurement output to provide the confidence needed going forward in the quality of postal service delivery not only in each country measured but also at on a global level.

As a proven and reputable audit services provider, PwC was pleased to support UPU with this challenge, leveraging our extensive experience in the postal industry, particularly in quality monitoring and auditing.

In agreement with the UPU Directorate of Postal Operations (DOP), we performed our activities for the two MSPs, UPU GMS and UNEX UPU TD, using the UPU Global Monitoring System Technical Design 3rd Edition (UPU GMS TD) with the following scope:

- Reperformance on samples of calculation of statistical design and allocation of links and items for the year 2022
- Reperformance on samples of item validation for the period from April to June 2022
- Verification of bundling on samples for the period April to June 2022
- Analysis of RFID diagnostic monitoring with a RFID service provider
- Analysis of Panel Management and planning and production of test items at the two service providers serving UPU GMS and UNEX UPU TD for GMS measurements.
- Quality of Service: Analysis of sent and received items for the months January to November 2022 for all countries in the UNEX UPU TD for GMS measurement
- Follow-up on recommendations in the 2021 report

1.2. Observations

Events like restrictions due to COVID-19 in some origin, destination and transit countries and the disruptions of international mail flow in and out of Ukraine and Russia, affected the measurement during 2022.

A major shortage of valid test items at UNEX UPU TD affected most of the measured countries, in particular during the first six months of 2022, with a partial recovery in the second half of the year. Traffic related to this event has been analysed in more detail due to the impact on the compliance to the technical design. This is considered a non-compliance issue in relation to the UPU GMS TD requirements.

The UPU Global Monitoring System Technical Design 3rd Edition, v1.0 (UPU GMS TD) released in August 2020 became effective from January 2022. The major impact noted is the decreased number of required valid test item volume.

Calculation of statistical design and allocation of links and items for the year 2022

Performed work	Result/observations	Impact for measurement
In close contact with key contact persons at the UPU International Bureau (UPU IB) and IPC, we performed a recalculation of the statistical design and allocation of links and items for a sample of two countries of the UNEX UPU TD measurement (a level A and a level B) and three countries of the UPU GMS measurement (a level A, a level B and	No deviations from the UPU GMS TD.	No impact.

Performed work	Result/observations	Impact for measurement
a level D), based on the rules of the UPU GMS TD.		

Item validation for the period from April to June 2022

Performed work	Result/observations	Impact for measurement
All items for the five DOs in the sample (the same as for the statistical design calculation reperformance) in the months of April, May and June 2022 were verified as being correctly marked as 'On-time' or 'Delayed'.	No deviations from the UPU GMS TD.	No impact.
The correct application of validation rules as stated in Appendix G of the UPU GMS TD (P1.1, P1.2, P1.5, P1.7) has been verified.	<p>Minor exceptions have been noted only for rule P1.1 (timely registration of dropping).</p> <p>For rule P1.4 (bundling), refer to the bundling paragraph.</p> <p>To be noted that for the IPC measured DOs in scope there are numerous test items not having an actual send date: refer to section 4.3.</p>	These exceptions have no impact on the measurement.

Verification of bundling on samples for the period April to June 2022

The same sample as in the reperformance validation has been also used for the verification of bundling

Performed work	Result/observations	Impact for measurement
Verification of bundling on the sending side based on planned date and on effective date.	<p>No or minor bundling observed.</p> <p>To be noted that for the IPC measured DOs in scope there are numerous items not having an actual send date: refer to section 4.2 (findings)</p>	We do not consider these deviations as non-compliance issues, but we suggest improving the controls: refer to section 4.4 (suggestion to relevant UPU bodies).
Verification of bundling on the incoming side.	<p>Bundling on the inbound side has been generally high for the countries in the sample, in some cases reaching 42% of the test items, exceeding the thresholds defined by the UPU GMS TD. This bundling is not generated by bundling on the sending side, which was minor.</p> <p>The reason is unclear, since the origin of the bundled items cannot be identified to specific countries. Bundling requests from designated operators could affect the measurement.</p>	<p>We do not consider this bundling a non-compliance issue, but we suggest the relevant UPU bodies to address the possible issue arising in future from the application of a systematic bundling check as defined in paragraph 18.1.7 of the UPU GMS TD.</p> <p>To be noted that it is responsibility of the designed operator to request a bundling check and that in future, in line with the UPU GMS TD automatic bundling check could be implemented by the Measurement Service Providers (MSP).</p>

Analysis of RFID diagnostic monitoring with a RFID service provider

Performed work	Result/observations	Impact for measurement
We performed an assessment of the processes and of the technologies deployed for the RFID with a service provider used by one of the QS link DOs.	No deviations from the UPU GMS TD.	No impact.

Analysis of Panel Management and planning and production of test items at the two service providers serving UPU GMS and UNEX UPU TD for GMS measurement.

Performed work	Result/observations	Impact for measurement
We performed an assessment of the Panel Management focussing on the planning and production of test items at the two service providers Quotas, servicing UPU GMS, and Kantar, servicing IPC UNEX UPU TD, for the GMS measurement. The analysis focussed on the recruitment, training and management of panellists, the test item production and test item circulation, as well as the data collection and validation.	<p>In this analysis, we identified major discrepancies in performance on panel management at Kantar compared to the previous years.</p> <p>No specific cause could be identified but several changes may have contributed to that:</p> <ol style="list-style-type: none"> 1) External global events making it difficult to keep panellists; 2) A change of the statistical design requiring changes in the distribution of panellists; 3) A shift from the Kantar internal solution used for the data collection and maintenance to a solution provided by IPC. 	Significant impact on the Q1 and Q2 'Valid on Target' numbers. For more details, refer on the next paragraph 'Quality of service: Analysis of sent and received items for the months January to November 2022 for all DOs in the IPC UNEX UPU TD for GMS measurement'.

Quality of Service: Analysis of sent and received items for the months January to November 2022 for all countries in the IPC UNEX UPU TD for GMS measurement

Performed work	Result/observations	Impact for measurement
Following up on the analysis of panel management for the countries with measurement operated by IPC we performed an analysis of sent items in relation to the expected allocation accordingly to UPU GMS TD, taking into consideration the implication on 'Valid on Target' (VOT) performance.	The results of this analysis, shown in detail in section 4.3, indicates that a widespread shortage of items occurred over several links to several destinations in the IPC UNEX UPU TD measurement for GMS. The shortage affected almost all measured countries in the first 6 months of 2022 with a recovery pattern at a later stage. For 13 out of the 15 measured countries the 'Valid on Target' reached is below 85%.	The statistical representativity of the measurement can be questioned, talking also in consideration additionally the asymmetry of the figures over the year despite the recovery occurred in the final months of the year.

Findings from the previous year

The two findings (Finding ID 1–2: Retention and training of panellists) from the 2018 audit, which only partially affect compliance and have low significance, are currently still open. These findings are related to conscious decisions made to improve operational processes that are not yet reflected in the current UPU GMS TD document. These findings should be considered together with the new point of attention for the relevant UPU bodies.

We note that the problem reported in 2018 for 'Test items not reaching destination' (Finding ID 3 in 2018) appears to have been at least partially remediated, with the average number of VOT from China in the months from August to November 2022 for the countries measured by IPC UNEX UPU TD for GMS being almost at the level of the overall average for all other origins.

The uncertainty remains about a specific measurement for one country in the UNEX UPU TD measurement that is struggling with the implementation of passive technology for reasons beyond the control of the operator.

General result

Based on our procedures as described in this report, a potential non-compliance issue was identified to the UPU GMS TD in relation to the shortage of items by UNEX UPU TD for the first six months of the year, for which compensatory measures need to be assessed.

Nothing else came to our attention that caused us to believe that the activities performed by UPU GMS, by UNEX UPU TD measurement systems or by the service providers in the audited areas were not compliant with the UPU GMS TD document.

This report has been prepared solely for the use of UPU in connection with the audit as requested by the UPU IB and should not be quoted in whole or in part without our prior written consent. No responsibility to any third party is accepted, as the report has not been prepared for and is not intended for any other purpose.

The procedures performed by us do not constitute either an audit or a review made in accordance with International Standards on Auditing or International Standards on Review Engagements. Consequently, we do not express any assurance on the information included in this report.

2. Scope of our work

The main objective of the external audit was to assess whether the methodology, its implementation and the calculation of quality of service (QS) measurement results by the two MSPs were compliant with the UPU GMS TD document in 2022.

The scope covered the following areas and components:

- Verification of the statistical design (incl. allocation of links and items)
- Recalculation of the Performance Measurement
- Recalculation of validation of item
- Verification of bundling
- Calculation and reporting of Quality-of-Service results
- Quotas – Panel Management and letter production audit
- Kantar – Panel Management and letter production audit
- IPC UNEX UPU TD – Panel Management audit
- UPU GMS – Panel Management audit
- Mieloo & Alexander – RFID Audit



3. Audit methodology and process

Based on our postal measurement experience, we have developed specific audit procedures that we applied in this engagement.

We performed an assessment of the current postal measurement procedures that will allow UPU to understand the quality of service they are getting from their service providers in comparison with what is required by the UPU GMS TD 3rd Edition document. We also provide clear insight on where improvements are needed and clear enforceable recommendations.

Our approach is:

- Independent
- Comprehensive
- Reliable and robust
- Statistically accurate
- Quality-driven and standardised
- Tested and proven over many years
- ISO 9001 consistent

While the methodology is standardised, PwC recognises that each client's environment and requirements are different. Hence, we customised it for this specific task, focussing on the four areas in respect of compliance to the UPU GMS TD document:

- Calculation and reporting of Quality-of-Service results
- Panel Management
- Quality control and validation
- RFID Diagnostic Monitoring

Our methodology this year was underpinned by the following tasks:

- Understanding the requirements of the UPU GMS TD specification document.
- Assessing the risks and mapping all elements in focus to our specific audit process (ref. diagram 1). We produced a viable, solid and efficient work plan.
- Collect information in appropriate mode: we know what should exist and how it can be assessed.
- Obtaining during the UPU and IPC interviews information and documentation by exchanging experience on postal measurement management with like-minded PwC people.
- Performing efficient walkthroughs with very experienced and skilled individuals of the key service supplier.
- Understanding deviations and confirming them with follow-ups. Performing recalculations wherever appropriate, leveraging our specific tools for this purpose.
- Formulating preliminary reports that can be validated.
- Producing a final report that is adequate for management and for those who have to work with it.
- Findings are formulated in a way that will help follow-up actions and improvements.

This methodology was used from the first year, confirming situation and progress, leveraging all of the experience from previous years.




Audit Process



Process execution against design at MSP and at organisation managing the systems

<p>Operations and IT set-up</p> <p>Panel set-up and organisation</p> <p>Quality assurance set-up</p>	<p>Data collection validation, organisation and transmission for implementation of statistical design</p> <p>Application of GMS technical design, especially geographical requirements, in implemented statistical design</p>	<p>Recruitment of panellists</p> <p>Panel performance management through KPIs</p> <p>Incentive management</p> <p>Panel training</p>	<p>Generation and preparation of test items</p> <p>Programming and integration of RFID tags</p> <p>Dispatch of test items</p>	<p>Test item circulation</p> <p>Registration of induction and delivery information and return of test items</p>	<p>Data entry and validation of panellist induction and delivery data</p> <p>Evaluation of panellist data</p> <p>Validation of panellist data against RFID data</p> <p>Diagnostic monitoring</p>	<p>Data analysis</p> <p>Exception reporting</p> <p>Proactive analysis to identify potential project risks</p> <p>Accuracy of calculations</p>	<p>KPI(s)</p> <p>Reporting according to timetable</p> <p>Recommendation from site survey process</p>	<p>Archiving of text items</p>	<p>Contingency planning</p> <p>Quality controls</p> <p>KPIs</p> <p>Change management process</p> <p>Process monitoring</p>
--	---	---	---	---	--	---	--	--------------------------------	--

Demonstrating understanding of GMS technical design by Measurement Service Partner (MSP)

 <p>Existence and extent of documentation for all audited areas</p>	 <p>Correct application of GMS technical design</p>	 <p>Implementation of country specific design parameters</p>	 <p>Implemented internal controls framework</p>
--	--	---	--

4. Audit results

4.1. Results per audited area

Based on our procedures as described in this report, an issue was identified concerning potential non-compliance with the UPU GSM TD in relation to the shortage of test items by UNEX UPU TD for the first six months of the year, for which compensatory measures need to be assessed.

Nothing else came to our attention that caused us to believe that the activities performed by UPU GSM, by UNEX UPU TD measurement systems or by the service providers in the audited areas were not compliant with the UPU GSM Technical Design 3rd Edition document.



The following table provides an overview of the results in the audited areas. When we noted at least one non-compliant finding, we have marked the area red; otherwise, it is marked yellow when there was at least one partially compliant finding. Areas are marked green when non-compliance issues were not detected in the given area. The numbers included in the table below indicate how many findings were identified per measurement area (in total 2, see detailed list in section 4.2




Measurement areas	UNEX UPU TD meas.	UNEX UPU TD meas. – PMC – Kantar	UPU GSM meas.	UPU GSM meas. –PMC– Quotas
A. Statistical design (sample design)				
B. System configuration and inputs				
C. Panel management				
D. Test Item production				
E. Test Item circulation (distribution/sending/receiving)				
F. Data collection, validation and processing				
G. Transit time calculations				
H. Statistical Analysis				
I. Reporting				
J. Archiving				
K. Quality Control				
L. RFID Diagnostic Monitoring system				

Compliance rating: Compliant Partially compliant Non-compliant

4.2. Detailed findings



The following list shows the current identified and open findings.

Finding ID	Area ID	Area description	Assessment area	Compliance	Issue description	Significance	Recommendation/Assessment results
1	C1	Panellists' recruitment questionnaires to ensure that UPU-specific recruitment requirements are satisfied	UNEX UPU TD measurement - PMC - Kantar	 Partially compliant	<p>Panellists' retention period</p> <p>The panellists were not informed, as part of the hiring process, about the requirement that they should be willing to participate for at least six months. This is not fully in accordance with section 7.2 of the UPU GMS TD document: "In all cases, panellists: [...] should be willing to participate for at least six months".</p> <p>However, we noted that the approach generally used to reduce the risk of not having the necessary number of panellists is not addressed by formally requesting the panellist to commit for at least six months but by having and managing back-up panellists.</p> <p>Finding remains open and unchanged in 2021.</p>	 Low	<p>We recommend either implementing a clause in the recruitment questionnaire to ensure the panellist is aware that he/she is expected to participate for at least six months or agreeing with UPU on updating the formulation of the UPU GMS TD.</p> <p>The UNEX UPU TD measurement system and Kantar do not fully agree with the recommendation as they express concerns because being formally bound by such a retention requirement may put off panellists from staying at least six months on the panel.</p> <p>Therefore, we suggest the UPU GMS measurement system and UNEX UPU TD measurement system formally agree on the next steps and assess whether the recommendation needs to be implemented or the formulation of the UPU GMS TD can be adjusted.</p>

Finding ID	Area ID	Area description	Assessment area	Compliance	Issue description	Significance	Recommendation/Assessment results
2	C6	Process of panellists' training	UNEX UPU TD measurement - PMC – Kantar	 Partially Compliant	<p>Training of panellists</p> <p>There was no formalised way to assess whether panellists have been sufficiently trained before starting to act as a panellist. However, we noted that the panellists' performance was monitored and that, in the case of low performance, the panellist was retrained.</p> <p>The UPU GMS TD document (section 7.3) mentions that “training should confirm that the panellist has understood the task involved and is able to carry it out as instructed”.</p> <p>In addition, the documented training programme for newly recruited panellists does not cover the topics on how to indicate the condition of the item received (envelope damaged, address label damaged or not fully legible, transponder missing, etc.). This is not fully in line with UPU GMS TD document (section 7.3.2) where it states, “instructions should indicate: [...] how to indicate the condition of the test item received (envelope damaged, address label damaged or not fully legible, transponder missing, etc.)”.</p> <p>Finding remains open and unchanged in 2021.</p>	 Low	<p>We recommend implementing an assessment process to ensure the knowledge of the panellist is tested before involving her/him as an active panellist.</p> <p>In addition, we recommend adding to the instructions provided to panellists a section on how to indicate the condition of the item received.</p> <p>The UNEX UPU TD measurement system and Kantar do not fully agree with the recommendation as they express concerns because they believe that training guidelines (via video, long-form written and FAQs) provide a comprehensive introduction to panellists' tasks. In addition, they monitor their panellists to confirm that they understand their duties. If deviations are observed, panellists will be retrained or dropped as appropriate.</p> <p>Therefore, we suggest the UPU GMS measurement system and to UNEX UPU TD measurement system formally agree on the next steps and assess whether the recommendation needs to be implemented or the formulation of the UPU GMS TD can be adjusted.</p>
3	-	-	UNEX UPU TD measurement		<p>Test items not reaching destination</p> <p>We noted that the UNEX UPU TD measurement system was affected by a large number of test items not reaching destination, even after a long period of time, despite being induced according to the UPU GMS TD document.</p>	 Medium	<p>It is recommended to monitor the situation.</p> <p>It is also recommended to make all possible efforts to resolve the situation for the operators having external difficulties introducing passive technology.</p>

Finding ID	Area ID	Area description	Assessment area	Compliance	Issue description	Significance	Recommendation/Assessment results
					<p>In particular, no test items at all induced between June 2018 and October 2018 reached the destination countries. Starting in November 2018, test items were registered again: 12 out of 1,178 in November 2018 and 33 out of 1,141 in December 2018. Please refer to section Update 2019 for the current situation.</p> <p>Since the test items were produced in line with the UPU GMS TD document and there are no indications they were not induced, this is not considered as a non-compliance matter, but the number of valid test items falling below the recommendations of the UPU GMS TD is influencing the performance measurement of the receiving countries. No similar pattern for the UPU GMS has been identified. The issue is known to UNEX UPU TD measurement and to UPU GMS, but no root cause has yet been identified.</p> <p>Update 2022 The root causes of the problem have been analysed in a pilot and, as a result, the UPU member DOs already implementing or joining the UPU QS link have been urged to make every effort to use passive technology for RFID. It has also been requested to increase the number of droppers from China to at least 2. The consequences of the problem have been addressed by introducing the re-weighting procedure. The finding on the test items not received appears to have been at least partially remediated, with the average number of VOT from China in the months from</p>		

Finding ID	Area ID	Area description	Assessment area	Compliance	Issue description	Significance	Recommendation/Assessment results
					August to November 2022 for the DOs measured in the UNEX UPU TD measurement being almost at the level of the overall average for all other origins. Complete resolution cannot be completely confirmed due to the circumstances described in finding #5. There is currently one designated operator in the UNEX UPU TD measurement still struggling with the implementation of passive technology for reasons beyond the control of the operator.		
4	F	Data collection, validation and processing requirements	UNEX UPU TD measurement		<p>We observed for the sampled countries some limited bundling on allocation. We also noted some low bundling caused by a delayed dropping of some panellists.</p> <p>To be noted that this bundling on the sending side is contributing to the much larger bundling seen on the inbound side (in some cases reaching 25% of the items), possibly driven by the COVID-related disruptions in the international logistics (refer to the points of attention for the relevant UPU bodies in section 4.4).</p> <p>Update 2022</p> <p>In our sample testing we noticed no bundling on allocation. There is some bundling caused by delayed dropping, but in this regard finding #5 is significant. Bundling on the receiving side continues to exist, reaching a value of up to 42% of the test items for a destination DO in the sample.</p>	🟡 Medium	<p>We suggest</p> <ol style="list-style-type: none"> 1) To further reduce the bundling in allocation. 2) To monitor bundling caused by deviations from droppers. <p>See also the points of attention for the relevant UPU bodies in section 4.4.</p>

Finding ID	Area ID	Area description	Assessment area	Compliance	Issue description	Significance	Recommendation/Assessment results
5	C	Panel Management	UNEX UPU TD measurement - PMC - Kantar	 Partially compliant	We observed a significant drop in performance in 2022 in the VOT KPI for the UNEX UPU TD measurement due to a very low number of test items sent and valid test items received. For a detailed analysis, refer to section 4.3. The main reason is insufficient recruitment and management of panellists.	 High	We recommend implementing more stringent quality control measures to ensure that the planning of the measurement achieves the defined KPIs. See also the points of attention for the relevant UPU bodies in section 4.4.

Compliance rating:  Compliant  Partially compliant  Non-compliant

Significance rating:  Low  Medium  High

4.3. Quality of Service: Analysis of sent and received test items for the months January to November 2022 for all countries in the IPC UNEX UPU TD for GMS measurement

4.3.1. Valid on Target over the year

The following analysis covers the 2022 allocation and the circulation from January to November 2022 of test items for all countries in the UNEX UPU TD measurement system, where issues in managing the panellists were observed. No significant issues for the UPU GMS were noted.

The base data used are all circulated test items from January to November 2022 (data were extracted around the end of December 2022, but the month of December has been excluded from the analysis to ensure a complete set of data over the analysis period). The correctness of the allocation calculation has been verified on a sample of countries by means of a recalculation.

Comparing Valid on Target (VOT, defined as valid/allocated accordingly to statistical design) per **destination** country shows that VOT was largely under the minimal requirements of the GMS UPU TD in the first part of the year, with a partial recovery at the end of the year.

Valid over Target												Overall
on Destination												62.6%
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	JAN-NOV
AT	16.7%	18.8%	28.1%	25.2%	40.9%	59.0%	77.5%	102.0%	116.2%	126.5%	123.3%	66.7%
CY	4.8%	18.2%	18.9%	13.7%	17.5%	13.0%	69.1%	111.1%	116.3%	112.6%	115.2%	55.5%
DE	28.5%	29.9%	48.4%	46.4%	48.6%	64.6%	76.0%	87.0%	81.3%	84.7%	90.2%	62.3%
DK	24.0%	24.0%	31.6%	45.6%	50.2%	49.2%	50.2%	79.2%	76.5%	89.8%	98.5%	56.3%
FI	35.5%	38.5%	55.8%	55.5%	60.6%	60.3%	72.5%	102.1%	108.0%	112.4%	105.8%	73.4%
HU	18.5%	18.5%	38.4%	42.4%	45.4%	39.5%	58.3%	89.0%	96.0%	90.1%	101.9%	58.0%
IE	8.9%	10.2%	36.7%	46.6%	57.7%	53.8%	56.4%	68.2%	89.2%	89.2%	84.9%	54.7%
IL	0.0%	0.0%	18.7%	26.1%	34.7%	37.7%	53.4%	92.6%	83.6%	93.7%	79.5%	47.3%
IS	41.6%	44.1%	65.5%	64.3%	61.8%	39.1%	45.4%	59.2%	108.4%	99.6%	88.2%	65.2%
IT	64.8%	69.9%	105.0%	95.3%	102.8%	89.8%	88.1%	111.4%	133.6%	133.2%	127.1%	101.9%
LU	43.8%	49.4%	62.1%	66.9%	71.4%	73.3%	75.5%	73.3%	95.0%	95.0%	95.4%	72.8%
LV	47.3%	48.4%	69.9%	77.3%	76.5%	78.8%	94.3%	143.9%	138.7%	131.3%	89.1%	90.5%
NL	38.8%	32.1%	35.5%	28.8%	33.4%	64.8%	73.0%	86.5%	100.0%	92.9%	108.2%	63.1%
SE	18.6%	14.6%	33.2%	26.2%	41.8%	51.5%	48.1%	60.8%	63.4%	84.7%	83.7%	47.9%
US	0.9%	1.9%	5.7%	10.4%	15.1%	17.2%	25.6%	45.1%	54.1%	52.4%	46.8%	25.0%

Legend

- over 85%
- between 60% and 85%
- under 60%

Fig 1: Valid on Target per destination country for the period January to November 2022

The recovery, which has been attempted with some boosting, cannot compensate the low level reached in the initial part of the year. At the end of November 2022, 13 out of 15 countries are under the minimum value of 85%.

It is possible, according to the UPU GMS TD rules, to perform a re-weighting for permanent links with a VOT under 85%. Since the correction has to be done for almost all measured DOs and due to the large asymmetry of the figure, it is doubtful whether these data from the UNEX UPU TD measurement system can be used to correctly reflect a reliable quality of service performance measurement. Decisions on compliance should be taken by the relevant bodies of the UPU.

4.3.2. Analysis on VOT of the origin side of the circulation

Comparing VOT per **origin** country shows that the reason cannot be attributed to a single or a limited number of sending countries. The analysis performed shows that VOT was under the minimum expected value of 85% for the large majority of the sending countries.

4.3.3. Analysis of planned dates for circulation test items vs target number of valid test items from the UPU GMS Technical Design

To help in identifying the reasons for the occurrence, we performed an analysis of how well the planning (based on the planned dates for the test items) matches the objective (the number of required target test items from statistical design). The values of planned vs target number of test items (Figure 3) show a generally regular pattern. The few exceptions can be explained by specific events (for example, difficulties getting Russian panellists to drop test items) and possible adjustments made over the year to compensate contingencies **This indicates that the test items have been planned appropriately.**

4.3.4. Analysis of effectively sent test items vs target number as per UPU GMS Technical Design

We noticed that several test items had no effective sending date registered. Some of these test items had an RFID registration on the incoming side. We considered that test items having a registration on the sending side or a RFID registration on the incoming side to have been effectively sent. We then compared the percentage of effectively sent vs targeted with a view to their origin. The analysis shows that the number of countries with insufficient results are fewer than that indicated by the analysis on VOT (section 4.3.2), but the number of originating countries causing the issue is not limited to a small number.

We then additionally calculated the percentage of test items with a RFID registration that were received, i.e. having a panellist registration date with a destination countries view (Figure 2). The figure shows that the insufficient registration of received test items is not limited to a few countries but distributed over a larger set of countries.

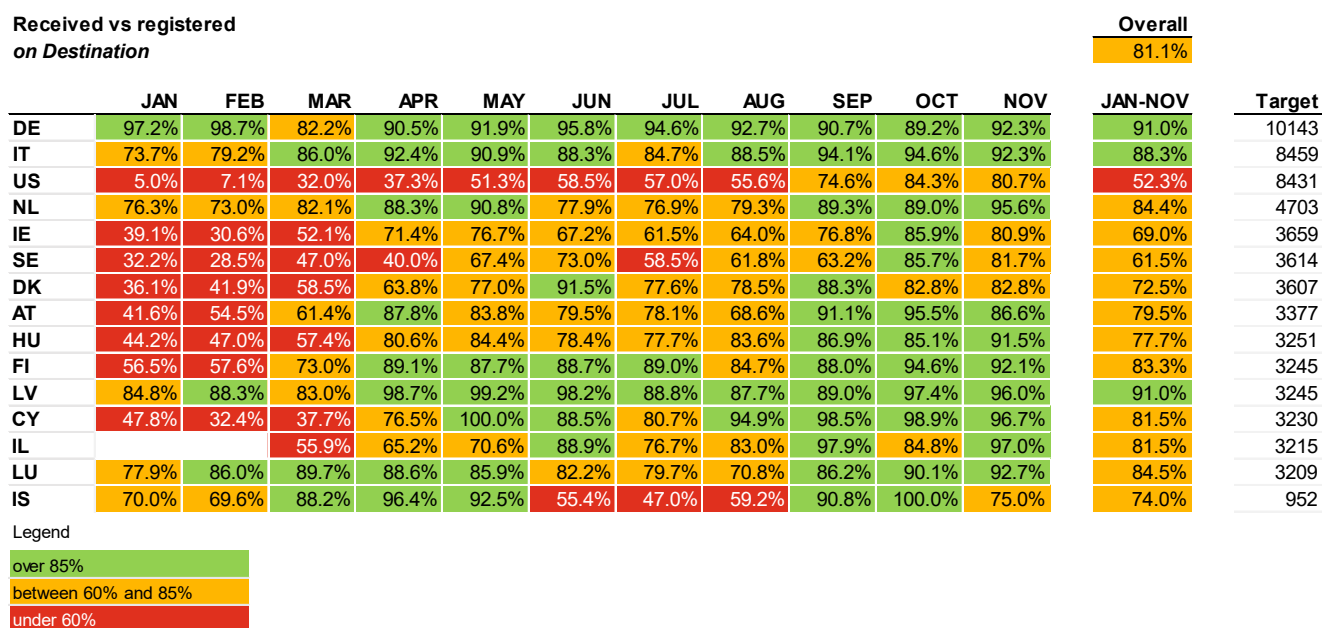


Figure 2: Percentage of the test items documented as received vs test items with an inbound RFID registration from a destination point of view

4.3.5. Analysis of test items that needed to be disregarded

We performed an analysis of the number of test items that had a RFID reading on the incoming side but have been disregarded for the measurement due to non-compliance of the data (marked as 'excluded'). Figure 3 shows that on the destination side the number of test items which needed to be excluded is relatively large. In some cases, the number of excluded test items is higher than the number of test items that are valid. Note that the two sets of test items ('valid' and 'excluded') are completely distinct; therefore, the proportion of the two sets combined may exceed 100%, since the proportion is **not** the number of valid test items that have been excluded.

Excluded vs valid on Destination												Overall
Country	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	Target
DE	1.7%	2.4%	15.2%	5.6%	13.6%	6.6%	11.4%	14.1%	16.6%	7.7%	7.0%	10143
IT	28.2%	22.1%	14.6%	8.6%	11.3%	13.0%	18.4%	15.7%	6.8%	3.1%	4.0%	8459
US	1800.0%	553.8%	162.5%	128.8%	107.5%	69.4%	65.6%	61.8%	25.8%	5.2%	1.8%	8431
NL	19.1%	31.7%	19.4%	10.6%	10.7%	26.8%	33.2%	34.8%	17.9%	10.7%	4.7%	4703
IE	129.6%	148.4%	72.3%	36.6%	28.4%	51.8%	56.4%	55.8%	23.5%	2.9%	1.2%	3659
SE	133.9%	175.0%	85.0%	94.9%	34.1%	24.5%	59.3%	50.8%	39.3%	4.7%	7.1%	3614
DK	140.3%	120.8%	88.4%	53.3%	35.1%	16.9%	32.5%	29.0%	15.2%	6.7%	4.1%	3607
AT	129.8%	88.7%	62.0%	25.4%	26.1%	29.5%	37.2%	47.4%	14.7%	5.3%	7.5%	3377
HU	116.0%	122.0%	69.2%	16.5%	24.4%	48.6%	40.5%	29.0%	15.8%	8.6%	2.5%	3251
FI	51.0%	53.8%	25.8%	12.7%	15.9%	14.7%	10.7%	14.1%	12.0%	4.3%	3.5%	3245
LV	14.8%	12.2%	20.6%	2.4%	2.9%	3.3%	8.2%	8.0%	3.7%	1.1%	2.1%	3245
CY	76.9%	85.7%	88.2%	24.3%	0.0%	8.6%	9.1%	2.7%	2.6%	2.3%	1.6%	3230
IL	0.0%	0.0%	42.0%	12.9%	12.9%	5.0%	39.2%	18.5%	2.7%	4.8%	1.4%	3215
LU	21.4%	12.1%	15.7%	17.3%	18.8%	21.4%	29.7%	45.9%	15.0%	7.5%	3.9%	3209
IS	18.2%	57.1%	9.6%	5.9%	20.4%	90.3%	94.4%	61.7%	4.7%	1.3%	0.0%	952

Legend

- under 10%
- between 10% and 20%
- over 20%

Figure 3: Number of test items excluded vs the number of the valid test items

There could be several reasons leading to an exclusion, in particular not correctly registered dropping or receiving dates, or no reception documented after a long time.

4.3.6. Conclusion

All the analyses show that the shortage of test items causing low VOT are not only referable to a specific link or country. The high number of test items that were not sent or receptions that were not documented indicates a problem with ensuring appropriate (available /trained) sending and receiving panellists in several countries.

The large number of test items excluded in several destination countries – despite having a RFID reading – reinforces the results of the analysis.

Other causes, for example the newly introduced possibility to send to a region in addition to a city, could eventually overlap for specific country or a link, but the issue appears to be of a general nature.

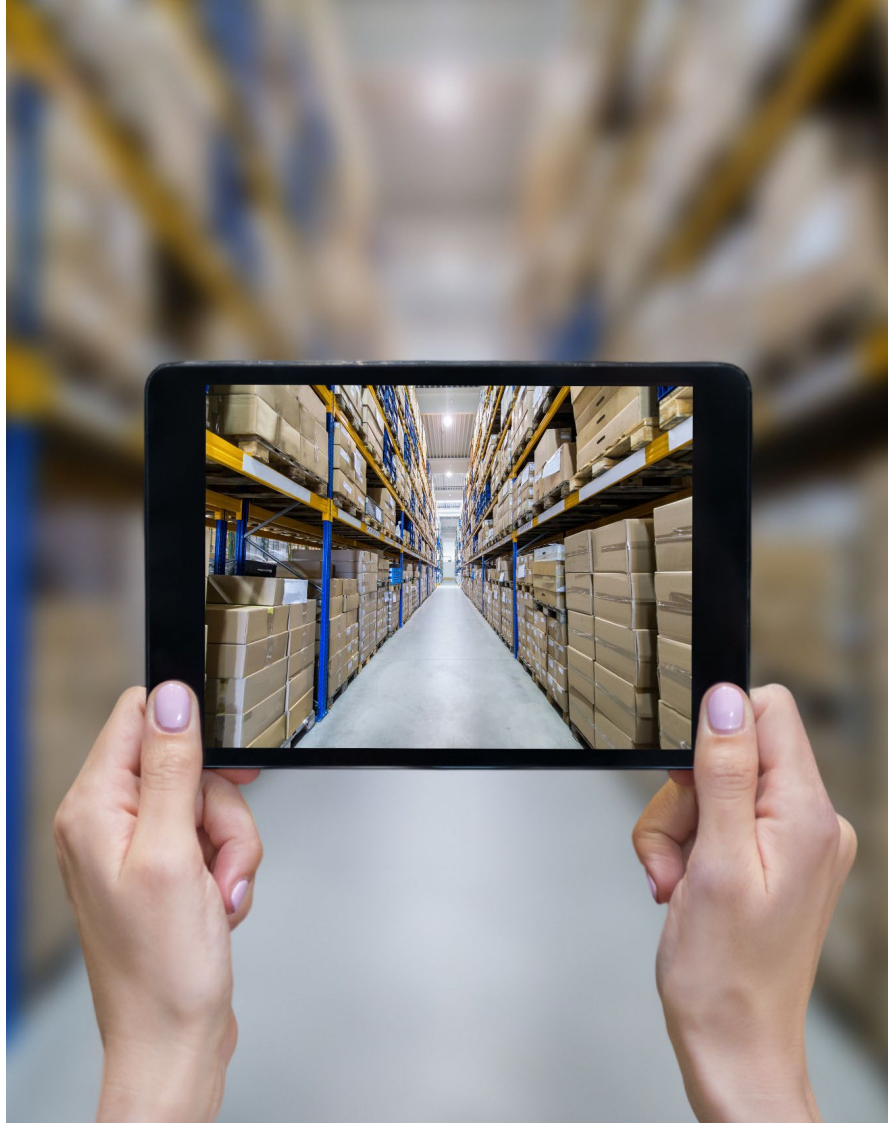
It should be noted that insufficiently managed changes affecting the organisations and the systems at the measurement service providers could cause a decline in Quality of Service that remains long undiscovered, leading to delays in the corrective measures and insufficient results.

4.4. Points of attention for the relevant UPU bodies

The following points for 2022 could currently represent a compliance issue and we suggest the relevant UPU bodies take a decision on how to handle this with regard to the performance measurement.

ID	Title	Description	Suggestion	Status in audit
1	Rules on handling bundling	<p>We noted that, for the countries in the sample, the number of test items that should be considered bundled according to UPU GMS TD (in paragraph 18.1.7 'Bundling on arrival due to operations') is relatively high, in some cases reaching 242% of the items. Since the bundling on the sending side is very limited, the bundling on the inbound side is very likely caused by irregularities in the international transportation.</p> <p>The rule in paragraph 18.1.7 of the UPU GMS TD states that test items identified as bundled on arrival due to operations will be excluded from the measurement. Identification can occur on request by the designated operators or on analysis of the MSPs. If automatic checks are introduced by the MSPs there is a potential of a large number of test items being excluded with a significant impact on VOT.</p>	<p>We suggest reconsidering the rule defined in paragraph 18.1.7 of the UPU GMS TD to avoid in future possible discussions in relation to the removal of large numbers of test items from the measurement.</p> <p>We also suggest for the performance measurement of 2023 that these test items be kept in the calculation to avoid reducing too much the VOT, with a larger impact on the measurement compared to the one caused by bundling.</p>	The utilisation of the bundled test items for the performance measurement is subject to the decision of the relevant UPU bodies.
2	Panel management	<p>We noted that insufficient management of the panellists can lead to a large reduction of valid test items and their asymmetry over the year, potentially causing the measurement to be non-compliant with the UPU GMS Technical Design.</p> <p>Insufficiently managed changes affecting the organisations and the systems at the measurement service providers could cause a decline in the Quality of Service that remains long undiscovered, leading to delays in the correction measures and insufficient results.</p>	<p>We suggest reassessing the approach used in monitoring/controlling the Quality of Service of panel management. In particular, deviations in VOT monitoring should be more proactive and timelier.</p> <p>Monitoring of panel management should be tied to the forecasting impact on VOT in order to have a better result on compensatory actions that might be effective only in the long term.</p>	The decision to utilise the available lower number of items for the performance measurement is subject to the decision of the relevant UPU bodies.

Annexes



A1 Rating criteria

Compliance rating criteria




The compliance rating indicated the compliance of the different assessment areas with the UPU GMS TD document.

Non-compliant means a clear violation of the UPU GMS TD document.

Partially compliant means a minor deviation from the UPU GMS TD document with no expected impact on the final measurement results. The significance rating provides indication on the severity and on the priority. Partial compliance can be related to

- a decision to deviate in order to improve quality in certain areas,
- a different interpretation of the UPU GMS TD document or
- a minor mistake in applying the rules.

Compliance rating:




-  Compliant
-  Partially compliant
-  Non-compliant

Significance rating criteria

The significance is an estimation of the impact on the measurement of the identified issue.

- Low means no impact on the measurement results.
- Medium means an impact on the measurement results that should be analysed, but expectation is that the impact does not change the measurement.
- High means that the measurement result is affected, and the implications should be analysed in detail.

Significance rating:

-  Low
-  Medium
-  High