



## **Towards improved quality of written patient records**

*Development and validation of language  
proficiency standards for writing for non-native  
English speaking health professionals*

**May 2017**

*Ute Knoch  
Catherine Elder  
Robyn Woodward-Kron  
Eleanor Flynn  
Elizabeth Manias  
Tim McNamara  
Barbara Zhang Ying  
Annemiek Huisman*



## ***Towards Improved Quality of Written Patient Records***

The *Towards Improved Quality of Written Patient Records* project was carried out with funding from the Australian Research Council Linkage Project LP130100171 and Cambridge Boxhill Language Assessment (Partner Organisation).

### **PhD scholars**

Sharon Yahalom

Simon Davidson

### **Associated Investigators**

Dr Alison Dwyer, Austin Health

**Acknowledgments:** The research team would like to thank the staff at Cambridge Boxhill Language Assessment, our partner organisation for their contributions to the project. We wish to thank the medical, nursing, and health information managers from the two hospital settings who participated in interviews and workshops in Phases 1 and 4, with particular acknowledgment of Dr Carolyn Kamenjarin and Ms Catherine Dooling at Goulburn Valley Health for their facilitation of data collection and their insights. Our thanks are also due to staff and doctors at Shepparton Medical Centre for assistance with the conduct of an on-site workshop, and to clinical academics in the Melbourne Medical School. We also wish to acknowledge the language assessors from Boxhill Language Assessment who participated in Phase 3 of the project, and Dr Susy Macqueen, Australian National University, formerly Language Testing Research Centre, University of Melbourne, for her input into Phases 1 and 2 of the project.

## TABLE OF CONTENTS

<b>0. Executive summary and recommendations</b>	<b>4</b>
<b>1. Introduction and Background to the project</b>	<b>7</b>
<b>2. The OET: A health-specific test of workplace communication</b>	<b>8</b>
<b>3. Study aims and research questions</b>	<b>9</b>
<b>4. Approach and training</b>	<b>10</b>
<b>5. Phase one: Understanding the writing processes that constitute written patient records</b>	<b>11</b>
5.1 Methodology	11
5.2 Findings	12
5.3 Summary of Phase 1	14
<b>6. Phase two: Establishing professionally relevant criteria</b>	<b>14</b>
6.1 Methodology	14
6.2 Findings	16
6.3 Summary of Phase Two	20
<b>7. Phase three: Developing a professionally-relevant rating scale for the OET writing task</b>	<b>20</b>
7.1 Scale criteria and top level descriptor development	20
7.2 Scale development workshop and pilot ratings - senior OET raters	21
7.3 Assessor training	22
7.4 Findings – OET raters application of professionally-relevant rating scale	23
7.5 Summary of Phase Three	26
<b>8. Phase four: Setting minimum standards for professional registration</b>	<b>26</b>
8.1 Methodology	26
8.2 Findings	28
8.3 Summary of Phase Four	31
<b>9. Recommendations</b>	<b>31</b>
9.1 Recommendations for Practice	31
9.2 Recommendations for Research	34
<b>10. References</b>	<b>36</b>
<b>Appendix 1: Rating scale used for large-scale trial with OET raters (Phase 3)</b>	<b>38</b>
<b>Appendix 2: Descriptions of scale criteria</b>	<b>39</b>

## **0. EXECUTIVE SUMMARY AND RECOMMENDATIONS**

This report sets out the background, aims, methods and findings of a project funded under an Australian Research Council Linkage Project grant with matching funding from Cambridge Boxhill Language Assessments, the partner organisation. The project was undertaken by a cross-disciplinary team consisting of language assessment specialists, health professionals and health professional educators, and experts in health communication. The project was conducted in four phases over a three and a half year period, and commenced in 2014.

The impetus for the project came from a concern that the current English language screening mechanisms used for overseas-trained health professionals may not adequately distinguish between health professionals with and without the requisite language skills to fully participate in English-speaking workplaces. This project focussed on written communication, following on from a previous project (Elder, McNamara, Woodward-Kron, Manias, McColl, Webb, 2013) which focussed on spoken communication.

The project focussed on the criteria used for the assessment of the writing sub-test of the Occupational English Test (OET), an English language screening test designed for twelve health care professions. The OET is used in Australia and other English speaking countries to assess the English language proficiency of overseas-trained applicants seeking registration in English-speaking countries. The task is currently assessed by language experts using a traditional linguistically-oriented rating scale.

The project had four aims:

- 1) to gain a detailed understanding of the writing practices that constitute medical records and to collect samples of written handover communication from real patient records (Phase 1).
- 2) to elicit the aspects of written communication valued by informants from three key professions (doctors, nurses, health information managers) and to translate these values into more professionally-relevant criteria for the OET writing task (Phase 2).
- 3) to involve OET language expert raters in the design of the new criteria and understand whether these criteria can be successfully implemented by the wider group of OET writing raters (Phase 3).
- 4) to empirically set new standards on the OET writing task by involving domain experts from medicine and nursing in a process called standard-setting (Phase 4).

The methodology used to address these aims is described in detail in the report which follows. The outcomes of the project's four phases are summarised below.

**Phase One:** During Phase 1, we gained a detailed understanding of how patient medical records are created, what aspects are included in patient records, who contributes to and accesses this documentation and what is valued in such records. The data showed that there were differences between a metropolitan hospital and a rural hospital in terms of how these medical records were created and stored, creating unique challenges at both institutions. We

were also provided with access to de-identified patient records from 199 patients from which written handover documentation was extracted as stimulus material for Phase 2.

**Phase Two:** During Phase 2, we elicited what doctors, nurses, quality and patient safety officers and health information managers value when reading written handover documentation. These values, or ‘indigenous assessment criteria’ (Jacoby & McNamara, 1999; Jacoby, 1998) formed the basis for a list of indicators which we created based on the interviews in Phase 1 and the workshop comments recorded in Phase 2. The indicators were grouped into (a) those which could be applied to the current OET test as well as (b) those which could either only be applied if the test was changed, or (c) if health professionals (domain experts) were involved in rating, or (d) those which could not be applied to a testing domain. The first group of indicators were then used as the basis for the rating scale developed and trialled in Phase 3.

**Phase Three:** During Phase 3, we went through several stages to develop a new, more professionally-relevant rating scale based on the list of indicators developed in Phase 2. We then trialled the new scale with fifteen experienced OET raters who each rated one hundred OET writing samples each. The results showed that the raters were generally able to apply the new criteria consistently, although it was clear from the qualitative data that more training was needed and that the raters would have liked more support from domain experts in the process.

**Phase Four:** In Phase 4, we convened standard-setting panels for both nursing and medicine to empirically set new standards on the OET writing sub-test, resulting in new cut-scores on the test. The results showed that if the new standards were implemented, fewer OET test takers would pass (i.e., receive an OET A or B band level) and that the medicine panel set more stringent passing standards compared with the nursing panel.

The results of the study have implications for both practice and research. These are outlined and explained in detail in the final section (Section 9) of this report and listed here.

With respect to the OET test design, and administration, we recommend that:

1. *the new professionally-relevant writing criteria (Appendix 1) are adopted for the OET writing sub-test,*
2. *the OET revise the existing specifications for the OET writing task to include scenarios and contextual information needed to elicit the qualities of communication valued by participants but currently not included in the writing task,*
3. *domain experts be involved in the training of OET writing raters and available for consultation during operational rating,*
4. *additional support materials be offered to raters during operational rating periods,*
5. *the OET provide training materials to accompany the new rating criteria to both test candidates and test preparation centres,*
6. *prior to implementation of any changes to the OET rating criteria, the test specifications or the test tasks, key professional groups be involved in a consultation process,*

## ***Towards Improved Quality of Written Patient Records***

7. *the new cut-scores be implemented for nursing and medicine.*

With respect to further research arising from this project, we recommend that:

8. *the new rating criteria be verified and trialled across a wider range of OET tasks for medicine and nursing and across tasks from all other ten professions,*

9. *additional standard-setting panels be convened to set cut-scores for the other professions.*

The findings from this project, as well as from a previous ARC Linkage study focussing on the spoken component of the Occupational English Test (LP0991153; Elder et al, 2013, if implemented, provide the opportunity of enhancing the Occupational English Test by making it more professionally-relevant.

## **1. INTRODUCTION AND BACKGROUND TO THE PROJECT**

Overseas trained health professionals play a crucial role in meeting Australia's health workforce shortages (Barton, Hawthorne, Singh & Little, 2003). International Medical Graduates (IMGs) will remain essential to primary, rural and acute health care delivery for the foreseeable future (Garling, 2008; Hawthorne, 2012). For example, in rural and remote Queensland, 46% of doctors are IMGs; in Victoria 36% of GPs are overseas trained (Hawthorne, 2012). Between 2001 and 2006, nearly 7000 internationally educated nurses (IENs) migrated to Australia (Hawthorne, 2012). Many IMGs and IENs are from developing countries, where English is not the official language (Mullan, 2005) and where clinical communication skills training tends not to be a foundational component of the medical or nursing curriculum (Dorgan, Lang, Floyd & Kemp, 2009). While there are local language and communication skills' interventions to enhance the quality and safety of IMGs and overseas trained nurses' spoken communication (Woodward-Kron, Stevens & Flynn 2011; Woodward-Kron, Fraser, Pill & Flynn, 2014; Konno, 2006), few interventions address the written communication skills required of health professionals.

Writing is used by healthcare professionals to communicate about patient care with other health professionals between departments, specialties and primary care. Written documentation provides a record of information used for monitoring patients' progress, for tracking their journey through the healthcare system, and for other administrative, legal and accounting purposes. The patient record includes a number of high level writing tasks: requests for investigations, referrals to another service to review the patient, and discharge and transfer summaries. Effective written records relating to clinical handover of patient information are crucial for safe and high quality patient care, and are particularly important in an ageing society where managing chronic conditions across multiple healthcare settings is a frequent experience (Manderson, McMurray & Piraino, 2012). Incomplete and inaccurate patient notes can lead to disrupted continuity of care as patients move between different environments and/or when various health professionals are involved in patients' care, leading to an increased risk of adverse events (Manias, Jorm & White, 2008).

One cause of incomplete and inaccurate patient records may be the language proficiency and written communicative competence of the health professionals contributing to the record. This highlights the need for defensible language tasks, criteria and standards for assessing overseas trained doctors' and nurses' readiness to communicate safely and effectively in the Australian healthcare setting. The need for fair and appropriate language proficiency measures emerged as a priority in stakeholder submissions to the 2012 parliamentary inquiry into the assessment and registration processes for International Medical Graduates (Standing Committee on Health and Ageing, 2012). The findings of this inquiry add urgency to the current investigation, which focuses on a particular proficiency measure designed expressly for use in the healthcare context: the Occupational English Test (OET).

## **2. THE OET: A HEALTH-SPECIFIC TEST OF WORKPLACE COMMUNICATION**

The Occupational English Test (OET) was designed to establish the adequacy of the workplace related language and communication skills of overseas-trained non-native speaking health professionals (McNamara, 1996). Administered by the OET Centre (partially owned by Box Hill Institute), the OET is administered in major cities across Australia and overseas. Doctors and nurses are the largest group of test takers.

The OET is a 4-skill test, measuring listening and reading comprehension as well as speaking and writing skills. While a recent study considered how performance is assessed on the speaking sub-test and proposed revisions to the associated rating criteria and minimum performance standards (Elder et al, 2013), there is scant available research on the writing component of the OET.

The OET writing task, which was the focus of the current study, is specific to each of the twelve health professions taking the test, requiring test takers to write a letter of referral based on written material (or case notes) provided to the test taker. The tasks are intended to reflect the writing demands of the respective professions. Candidate performances on these writing tasks are used to make inferences about their writing abilities in non-test situations, such as in healthcare settings, and on a range of writing tasks relevant to their workplace. However, test performances on the OET writing task are currently rated by linguistically-trained judges against a set of linguistic criteria developed by McNamara (1996). There is no routine involvement by health professionals in the judging stage; this absence of health professional input has implications for the relevance of the task for assessing candidates' communicative readiness for the workplace.

In order to address this issue, Jacoby and McNamara (1999) propose the use of criteria indigenous to the specific communicative context, which specific purpose tests like the OET, are designed to target. Identifying such criteria entails sustained engagement with the relevant real world context in order to find out what aspects of communication are valued by those involved. The claim is that identifying valued communicative practices and basing assessment criteria on these valued behaviours will yield more valid and professionally relevant measures of communicative competence.

The current study therefore sets out to gain an understanding of the health professional perspective on written communication in a range of healthcare settings. The insights thus gained were used to inform subsequent revisions to the current OET criteria, against which a candidate's performance is assessed. Health professionals were also involved in setting minimum standards of performance on the test, to ensure that those admitted to the profession were able to communicate safely and effectively in the workplace. The specific aims and research questions that guided the study are set out in further detail below.

### 3. STUDY AIMS AND RESEARCH QUESTIONS

The *first* aim of the study was to gain an understanding of the writing practices that constitute patient records, the types of documents included in patient records that are written or read by nurses and medical professionals and used as handover documents at different transition points of care, the purposes of different documents as well as the structure and layout and the type of information included in these documents. Because health information managers (i.e., those responsible for reading and coding health records for funding purposes) are a key stakeholder group when it comes to patient records, we also included members from this group in the Phase 1 interviews.

The *second* aim of the study was to gain a detailed understanding of the aspects of written handover communication valued by both nursing and medical professionals with the aim of eliciting indigenous assessment criteria (Jacoby, 1998; Jacoby & McNamara, 1999); that is, the types of criteria that professionals draw on when engaging with these types of documents. These criteria were elicited by showing domain experts (nurses and doctors) samples of written handover documents, in particular, referral letters and discharge summaries extracted from real patient records. Based on these values, more professionally-relevant criteria were created for the Occupational English Test.

The *third* aim of the study was to train experienced raters from the Occupational English test in the use of these professionally-relevant rating criteria and to evaluate how well the new rating scale functioned both quantitatively (via statistical analysis of scoring outcomes) and qualitatively (by eliciting feedback from the rater participants).

The *fourth* aim of the study was to set new passing standards for entry into the nursing and medical profession by involving groups of domain experts from the two professions in a process called standard-setting. This process required doctors and nurses to review sample writing performances from the OET test and judge the readiness of each writer for entry into the profession.

The following research questions were addressed in the study:

- 1) What are the practices that underlie the creation of written patient records?
- 2) What aspects of written communication do health professionals (nurses, doctors) value?
- 3) Can such professionally relevant criteria be used as the basis for language assessments carried out by *language experts* of migrant health professionals seeking registration in Australia?
- 4) What *minimum standards* should be set for professional registration of migrant health professionals?

#### 4. APPROACH AND TRAINING

The research questions listed above were addressed over a three and a half year period by a multi-disciplinary research team including experts in language testing, applied linguistics, medical communication, and medical and nursing educators. The project also provided the opportunity for two PhD students to work on data related to the project, one funded by the Linkage grant.

The project was conducted in four phases as summarised below in Table 1. More details of each phase and its outcomes are provided in the following sections of the report.

**Table 1.** Project phases and timeline

<b>Phase</b>	<b>Timeline</b>	<b>Aim</b>	<b>Outcomes</b>
<b>ONE</b> (Research Question 1&2)	September 2014 – June 2015	Gain detailed understanding of the writing processes that constitute written patient records & collection of 200 written patient records	<ul style="list-style-type: none"> <li>• Understanding of writing processes</li> <li>• Written documentation (referral letters and discharge summaries) extracted from written patient records</li> </ul>
<b>TWO</b> (Research Question 2)	April 2015 – December 2015	Establish professionally relevant criteria for assessing clinical communication skills of non-native speakers of English	<ul style="list-style-type: none"> <li>• Indigenous criteria representing what doctors and nurses value in written communication</li> </ul>
<b>THREE</b> (Research Question 3)	January 2016- December 2016	Development and trialling of new rating criteria for the OET writing sub-test	<ul style="list-style-type: none"> <li>• New criteria for the OET writing sub-test based on the outcomes of Phases One and Two.</li> <li>• A framework for creating professionally-relevant rating criteria for specific purpose language tests</li> </ul>
<b>FOUR</b> (Research Question 4)	June 2016- December 2016	Set minimum standards on the OET writing sub-test for professional registration of migrant nurses and doctors	<ul style="list-style-type: none"> <li>• Minimum passing standards set by health professional using the existing and new criteria.</li> </ul>
<b>WRAP-UP</b>	January 2017 – June 2017	Report findings	<ul style="list-style-type: none"> <li>• Final report.</li> </ul>

## **5. PHASE ONE: UNDERSTANDING THE WRITING PROCESSES THAT CONSTITUTE WRITTEN PATIENT RECORDS**

### **5.1 Methodology**

Two large Victorian hospitals agreed to participate in the first two phases of the study, one metropolitan and one rural. They differed in their practice in relation to patient records in that while both used hand written hard copy records to record contemporaneous inpatient events, one used electronic discharge summaries and letters to referrers while the other used a mixture of typed and hand written handover documents which were scanned following the discharge of a patient.

Phase One comprised interviews with key stakeholders in both hospitals in relation to patient records. Semi-structured interviews were conducted with doctors (N=18; mean years of experience 14), nurses (N=31; mean years of experience 25) and health information systems staff (N=6; mean years of experience 15). Health information systems staff were not involved as focus group participants in the project because they are not a target test taker group of the OET. However, they are important stakeholders as they are responsible for coding the patient records for hospital funding and therefore key readers of written hospital documentation.

The interviews focussed on a range of topics, including the types of documents participants read and contributed to, the intended readership of each of these documents, how key handover documents were created and for what purpose, the structure of the documents, what information needed to be included and what made for a good and poor document. In particular, we focussed on discharge summaries and referral letters as these are the documents closest to the writing task used in the OET.

The interview data was transcribed and then analysed qualitatively by coding for salient themes and sub-themes. Coding was undertaken by both applied linguistics and medical educator project staff to ensure interpretations were as accurate as possible. A sub-sample of interviews were also double-coded and the inter-coder reliability was 0.87 (Cohen's Kappa).

The findings from the interviews are presented below. As the interviews also informed the second research questions, any codes that related to what participants value in the handover documents are described with the findings of Phase 2 below.

A further aim of the first phase was to collect full patient records from the two hospitals. These were all from patients who were discharged within a twelve month period from the year preceding the commencement of data collection. Records were sought from weekday as well as weekend discharges from medical and surgical wards. The aim was to collect between 90-100 patient records from each hospital and from a diverse sample of clinical settings. The records were then redacted for any identifying patient, hospital, and clinician information. The redacted records were then used to extract the stimulus materials (referral letters and discharge summaries) to be used in the workshops in Phase 2 (described below).

## **5.2 Findings**

The findings for the writing practices that contributed to discharge summaries and referral letters in the two hospitals were similar in regards to their purpose, audience, contributors, and components; however, they differed in the participating hospitals' use of digital platforms, digitisation of the records and templates to record the admission and transferral of patients, the degree of auto-populated text and free text, and in what form and method of transfer the record reached its intended audience. There were also reported differences in practices within hospitals across units.

*Discharge summaries:* The nursing and doctor participants reported the primary purpose of the discharge summary was summarising for general practitioners and other external agencies what happened during the patient's stay while in hospital. Health information service participants emphasised that the discharge summary was the primary document for coding purposes as it provided the principal diagnosis, allowing the hospital to charge for clinical services provided. Discharge summaries were reportedly written almost exclusively by interns (doctors in their first year post graduation). Information that needed to be included in the discharge summary was:

- patient demographic details, admission and discharge dates;
- principal diagnosis, presenting complaint, associated conditions;
- complications, progress, management, investigations;
- medication changes;
- discharge plan, and follow up;
- information about who completed the summary, designation, and date.

While one hospital had implemented an electronic health record, the other hospital was in the process of moving to a digital record. In this process, some discharge summaries were hand written then typed up, with the medical record then scanned as a PDF.

General concerns about discharge summaries were that they were often incomplete, resulting in the need for follow up and subsequent time inefficiencies for general practices and health information services staff. Reasons for incompleteness or missing discharge summaries were seen to be a lack of understanding of the dual purpose of the discharge summary and that they were written by the most junior of doctors with little education or monitoring from more senior medical colleagues.

*Referral letters:* Medical records often include referral letters from the hospital to outside health professionals as well as referral for patients coming into hospital. Referral letters can be faxed or emailed or the patient can also present the referral letter in hard copy in person. Outgoing referrals from the hospital tend to be faxed while within the hospital, electronic referral systems are in place for between unit referral and verbal referrals were also reported.

## *Towards Improved Quality of Written Patient Records*

The nursing and doctor participants reported that the primary purpose of referral letters was twofold: summarizing the medical treatment so far, and specifying what needs to be done by the next health professional. According to participants, referral letters should therefore include:

- patient demographic details;
- relevant past medical history;
- current condition(s);
- relevant previous results and investigations related to the current condition;
- patient management;
- pending test results;
- treatment plan, what help is requested.

Participants reported that the majority of incoming referral letters were written by General Practitioners. When coming from the hospital, they were usually written by the treating doctor or specialist. Occasionally they were written by more junior doctors. The intended audience of referral letters was generally doctors or specialists, but referral letters were read by a wider audience, including nurses who cared for the patient during the stay in their ward.

General concerns about referral letters were that they were either not specific enough, for example in their request, or that too many irrelevant details were included. Both cases resulted in having to spend time assessing the patient to establish the problem and to make a treatment plan, which in fact was reported as one of the purposes of the referral letter.

Although the focus of phase 1 was to gain a greater understanding of the writing practices contributing to discharge summaries and referral letters, the following quote from a health information services manager highlights the importance of the medical record as a communication tool, as a patient safety and quality of care mechanism, and as a tool that has important budgetary and planning implications for the hospital. The multiplicity of roles and overall importance of the medical record identified by the informant underscores the need for the OET writing task to reflect the values and purpose of what constitutes effective written communication from multiple stakeholder perspectives.

Going back to health information 101, the record is a tool for communication [...] but it's also central to what the hospital does. Without that record you can't treat the patient effectively next time they come in. Without the record there is no money for funding, the position to treat that patient. Without that record, there is no data for building the new hospital that's required or the new theatre that's required, so it's very much a communication tool but it's a business communication, it's a treatment communication, it's a planning communication tool, so there's many uses for it, that the clinician at the bed side may well not understand, but if they write effectively for treating the patient, the other things will follow well too. (HIS)

Finally, phase 1 also involved the collection and redaction of records from both sites. In total, 199 records were collected. The medical records ranged in scope from a small number of pages to voluminous documents. We identified approximately fifty different component documents in the medical record, including an admission risk screen, nursing care plan,

prescriptions, ED to ward transfer checklist, mobility assessment, anaesthetic chart, as well as investigations and observation notes. The inclusion of these documents would lend authenticity to the OET writing task and provide important background information for the discharge summary and referral letter writing tasks; however, the extent and diversity of these accompanying documents preclude their inclusion in the OET writing task as currently implemented and assessed.

### **5.3 Summary of Phase 1**

The purpose of Phase 1 was to provide background information on the writing practices associated with the medical record, with a particular focus on discharge summaries and referral letters as these constitute the writing task of the OET. These practices included identifying the format and pathways of these documents, who contributed to them and who received them, as well as identifying their purpose. Informants were clinicians from a range of specialties, some of whom had educator roles, nurses, and health information service staff. Informants spoke in general about these aspects, including what they valued as effective written communication, which is the focus of the next section. Further, phase 1 included retrieving 199 medical records in total from both hospital sites, whose content assisted data collection in Phase 2.

## **6. PHASE TWO: ESTABLISHING PROFESSIONALLY RELEVANT CRITERIA**

### **6.1 Methodology**

Phase Two of the project investigated the values of doctors, nurses and health information managers when reviewing samples of written documentation, in particular discharge summaries and referral letters. These documents were extracted from the patient records from the two hospitals. After careful review of all the referral letters and discharge summaries relating to the last admission of each patient, ten documents were chosen. These documents were selected to represent a range of writers and audiences, a range of medical conditions and diagnoses and a range of features related to writing or language (e.g. unusual organisation of the information in the document, letters written by non-native speakers with grammatical inaccuracies). We included some handwritten documents and ensured that we represented documents from both the rural and the metropolitan setting. These were slightly different for nurses and doctors, with some overlap, while health information managers reviewed the same documents as doctors. If the document was originally written by someone in the hospital (rather than a GP), the participants also had access to the complete (redacted) patient records.

Thirty-one nurses, eighteen doctors and six health information managers took part in a series of small workshops. Nurses were drawn from a range of sub-disciplines and work areas (e.g. nursing unit managers, clinical nurse specialists, district nurses etc) and had an average experience working in their profession of 25 years. The eighteen doctors who participated were also drawn from a range of disciplines (e.g. general practitioners, ICU senior staff, an

## *Towards Improved Quality of Written Patient Records*

anaesthetist, oncologists, registrars etc.) and had worked in their profession for an average of 14 years. The six health information managers had an average of 15 years of experience.

Workshops for each of these professions were held separately. Following a general introduction and overview of the project, participants worked in pairs or groups of three with one facilitator. Participants were shown one document at a time and asked to comment on the strengths and weaknesses of each. The discussion was therefore not constrained to language and writing issues, to allow the participants to verbalise their indigenous criteria. The discussion about each document was recorded, transcribed and then analysed for themes and sub-themes by several team members together to ensure that applied linguistic and language testing staff understood the data fully. Through a process of refining the draft scheme and applying the new scheme again to all data, the final coding scheme was developed.

The codes identified in the data focussed on aspects relating to the text itself (which was the key interest of the study) but also focussed on the conditions the document was presumably created under (e.g. whether it was written at the side of a patient bed; or under time pressure), when it was written, timing of the referral, effects the letter may have (both bad and good) and professional judgements of the writers.

As the purpose of the study was to identify indigenous criteria which could be applied to the rating criteria of the OET, any quotes that related to textual features were then extracted from the data set. Values expressed by the health professionals in relation to particular documents were converted into a list of general indicators of document quality. Below are two examples of how we moved from workshop extracts to indicators. Extract 1 below, from Doctor workshop 2, which involves two GPs discussing a referral letter.

P: I find it quite difficult sometimes when you are referring to a consultant, I don't like to be saying 'Well this is what I would like' because sometimes consultants don't like being told what to do

(Doctor workshop 2, Document R13, Participants 11&12 - GPs)

### **Extract 1**

Similar extracts resulted in the indicator: 'Document appropriately reflects differences in position and clinical discipline between writer and reader'.

Extract 2 was obtained from another doctor workshop, where two participants are reviewing a referral letter.

P: We do jump around a little bit

P: yes, well you get the presenting complaint and then some past history

P: this I found a bit confused about the pregnancy and then the sister, the order of that...

[...]

P: Organisation of it, yeah, I think too

[Doctor workshop 1, document M12, Participants 6&7]

### **Extract 2**

Similar extracts to those of extract 2 resulted in the following indicator: structure/organisation of document is logical.

Based on this procedure, we created a list of 48 indicators, which were then grouped into several sub-areas, as can be seen in the following section. All 48 indicators were carefully reviewed by the project team to ensure their wording captured the qualitative data as closely as possible.

## 6.2 Findings

Table 2 presents the full list of 48 indicators created based on the workshop data as well as the sections of the Phase 1 interviews relating to the textual features of the handover documents. These are presented under some broad groupings (or constructs), which we created in an initial step towards rating scale design. It is important to note, however, that several of the indicators could have been grouped under more than one heading (or construct) and that, as a group of researchers, we decided collectively where these would be best placed for this purpose.

**Table 2.** List of indicators derived from workshops and interviews

<b>Constructs and checklist indicators</b>
<b>Content &amp; Audience awareness</b>
Language and content is appropriate to the specified/intended audience(s)
Language and content is appropriate considering the shared knowledge between writer and recipient (i.e. doesn't include what both know but explains things other person needs to know)
Content addresses what reader needs to know to continue care
Document effectively communicates doubt about need for referral or about symptoms reported by patient
Document effectively communicates uncertainty about diagnosis or future management
Content and style is appropriate to multiple (intended) audiences and future possible uses of document
Document appropriately reflects differences in position and clinical discipline between writer and reader
<b>Purpose</b>
Document achieves its purpose/ is useful/effective
Purpose of document is clear
Document fulfils multiple purposes (e.g. informing GP, note to self; informing patient or family)
Document achieves function/purpose of increasing health literacy of patient or patient's family
<b>Genre/Style</b>
Writing/style is appropriate to genre of document
Writing is typical of discipline of writer
Document is professional (formal) and clinical (factual) (not emotional or chatty) and objective
Document is appropriate to health care disciplines and level of knowledge of recipient
<b>Conciseness &amp; Clarity</b>
Document is concise/succinct; length and level of detail is appropriate (to case) and reader
Information is presented clearly
Information summarizes patient's stay/case effectively
<b>Organisation</b>
Structure/organisation is appropriate and logical (coherent) (to document) (might follow chronological timeline of events)

## Towards Improved Quality of Written Patient Records

Document is prioritizing/highlighting key information for reader (and/or visually/typographically highlighting important information)
Document follows well-known structure (such as ISBAR, SOAP <sup>1</sup> )
Sub-sections of document (e.g. medication lists, results) are well organised
Structure of document is highlighted to reader by e.g. sub-headings, bullet-points, paragraphing
<b>Sufficiency/completeness/Relevance of content</b>
All key information is included/no key information is missing
No unnecessary/irrelevant information is included
Information is sufficient for reader (including coder)
Writer has done more than just produce generic letter
<b>Professionalism and clinical ability of writer</b>
Document shows effort and attention to detail
Document shows experience and understanding of the referral process
Writer shows patient centred approach (has got to know patient well, has relationship/rapport with patient; has respect for patient (does not judge patient); and has involved patient in decision-making)
[in case of referral letters] There is a sense that the referral is genuine and justified and not a handball
Writer is displaying collegiality and professional courtesy
Quality of the document suggests clinical competence of the writer
(if also read by patient), statements in document will not create expectations (from patient) that next health professional cannot deliver
<b>Layout/Presentation</b>
Document is legible [in handwritten documents]
Bullet-points and prose are used appropriately
Document is tidy and well laid out
Paragraphing is used effectively and logically
[for electronic letters only] Fonts and font sizes are consistent
<b>Language</b>
Technical language and abbreviations are appropriate for the recipient and the document
Polite language is used appropriately (e.g. by hedging requests)
Spelling and capitalisation is accurate
Grammar is accurate
Sentence structure (whether short sentence or sentence fragments) chosen is consistent and allows the language to flow
Language used is respectful of patient
<b>Accuracy of content</b>
Content is accurate
Medication lists and dosages are accurate
<b>Timing</b>
Document was written immediately at discharge (not later) [documents coming from hospital only]

A further list of 26 indicators related to the content items that could be expected in a discharge summary or referral letter (Table 3). We have listed these separately, as these are more useful in the stage of creating test specifications, rather than applicable to a rating scale, where not all of these might be relevant to each task. They are, however, a useful list of points that could be supplied to raters (with a selection of these shown as relevant to a particular task) to ensure key information is present in an answer.

**Table 3.** Content items mentioned in workshops and interviews

<sup>1</sup> ISBAR: Identify, Situation, Background, Assessment and Recommendation; SOAP: Subjective data, Objective Data, Assessment, Plan

## Towards Improved Quality of Written Patient Records

Content
Key patient details (name, gender, address, GP, relevant recipients) are included
It is clear who has written document (including name, designation and contact details)
Key dates are listed and clear (DOB, practice attendance, past and planned examinations, discharge, date when document was compiled)
Principal diagnosis is stated clearly (where possible, even if provisional) and differential diagnoses are described
It is clear which diagnoses are active and which are dormant (if applicable)
Symptoms or presenting complaint are described
Clear context to case is set out
Specific request or question is posed (and this is reasonable/realistic)
Concern of writer is specified (if applicable)
Urgency has been assessed and is clear from the letter (if applicable)
Relevant past medical history is sufficiently described
Social/family history is sufficiently described (if applicable)
Patient level of independence is specified (if applicable)
Patient consent and life choices are clearly flagged (if applicable)
Any food restrictions or food allergies are described (if applicable)
Relevant/key investigations are detailed (if applicable)
Findings/results are described
Present or immediate past management of patient is described
Current medications are listed (e.g. medications and doses at discharge; at time of referral) or NIL
Recent medication changes are highlighted and unambiguous with explanations where necessary (if applicable)
Potential allergies to analgesia or medications are provided (if applicable)
Changes following or reactions to medications are described (if applicable)
[discharge summary] Patient's stay is described (including changes/complications/response to treatment/effectiveness of treatment)
A clinical synopsis is presented
Follow-up/treatment plans and responsibilities are clear (including instructions to family/parents)
[discharge only] Discharge destination is clear (if applicable)

The 48 indicators listed in Table 2 could not all be applied as formulated to the OET writing task. Many indicators were applicable to the OET task as it is currently designed, but many others are not relevant to the current specifications. For this reason, the multi-disciplinary project team met to code each indicator into one of the following categories:

- A – applicable to current OET tasks
- B – applicable to OET tasks if the current specifications are changed
- C – not suitable to be judged by language-trained raters
- D – not suitable to testing domain

Table 4 below presents all the indicators which were placed into Category A (applicable to current OET task).

**Table 4.** Indicators relevant to current OET task

<b>Constructs and checklist indicators</b>
<b>Content &amp; Audience awareness</b>

## *Towards Improved Quality of Written Patient Records*

Language and content is appropriate to the specified/intended audience(s)
Content addresses what reader needs to know to continue care
<b>Purpose</b>
Document achieves its purpose/ is useful/effective
Purpose of document is clear
<b>Genre/Style</b>
Writing/style is appropriate to genre of document
Document is professional (formal) and clinical (factual) (not emotional or chatty) and objective
Document is appropriate to health care disciplines and level of knowledge of recipient
<b>Conciseness &amp; Clarity</b>
Document is concise/succinct; length and level of detail is appropriate (to case) and reader
Information is presented clearly
Information summarizes patient's stay/case effectively
<b>Organisation</b>
Structure/organisation is appropriate and logical (coherent) (to document) (might follow chronological timeline of events)
Document is prioritizing/highlighting key information for reader (and/or visually/typographically highlighting important information)
Sub-sections of document (e.g. medication lists, results) are well organised
Structure of document is highlighted to reader by e.g. sub-headings, bullet-points, paragraphing
<b>Sufficiency/completeness/Relevance of content</b>
All key information is included/no key information is missing
No unnecessary/irrelevant information is included
Information is sufficient for reader (including coder)
<b>Layout/Presentation</b>
Document is legible [in handwritten documents]
Document is tidy and well laid out
Paragraphing is used effectively and logically
<b>Language</b>
Technical language and abbreviations are appropriate for the recipient and the document
Polite language is used appropriately (e.g. by hedging requests)
Spelling and capitalisation is accurate
Grammar is accurate
Sentence structure chosen is consistent and allows the language to flow
<b>Accuracy of content</b>
Content is accurate
Medication lists and dosages are accurate

Table 4 lists the 27 indicators which apply to the current OET task. The complete construct 'Professionalism and clinical ability of writer' was deleted because it was not possible to operationalise under the current task specifications. A number of indicators in this section, as seen in Table 2, relate to concepts of using a patient-centred approach and are probably the types of indicators that would be helpful to include in the OET to generate positive washback on candidates and their teachers in preparing for the test. We have included some recommendation in relation to this later in the report.

As this list of 27 indicators by themselves are not practical to apply in a proficiency test setting, further work was necessary to create a rating scale that could be applied by the language-trained raters used by the OET. These steps are described further under Phase 3.

### **6.3 Summary of Phase Two**

In Phase Two we elicited the aspects that key groups of stakeholders (doctors, nurses, health information managers) reading and contributing to patient records value in written handover communication. This was elicited by providing the health professionals with stimulus materials extracted from patient records provided by two hospitals. The “indigenous criteria” underlying health professionals’ feedback on these handover documents were uncovered via a thematic analysis. The themes uncovered were converted to a list of 48 indicators. Since not all of these indicators could directly be applied to the current OET task, the list was shortened to 27 indicators. The process of converting these indicators into a rating scale for the OET, is described in the following section.

## **7. PHASE THREE: DEVELOPING A PROFESSIONALLY-RELEVANT RATING SCALE FOR THE OET WRITING TASK**

The research question addressed in Phase Three was “Can such professionally relevant criteria be used as the basis for language assessments carried out by *language experts* of migrant health professionals seeking registration in Australia?” To answer this research question, the indicators developed in Phase 2 needed to be converted into a rating instrument that is practical for language-trained OET raters and then trialled. This was done in several stages, which will be outlined in this section.

### **7.1 Scale criteria and top level descriptor development**

The project team carefully reviewed the 27 indicators prepared in Phase 2 and grouped these into the following six criteria:

- **Purpose:** The workshop and interview participants very frequently mentioned that the purpose of a document needs to be immediately identifiable. For this reason, this was given its own category, rather than adding this into the content category. Identifying the purpose gives health professionals a quick and precise sense of what is asked of them.
- **Content:** This criterion was designed to focus the raters’ attention on whether the key information is included in the document (i.e. is everything that is needed to continue care present) and whether the information present is accurate. This criterion taps into audience awareness in that the writer needs to be aware what information is needed by the reader.
- **Conciseness and clarity:** This criterion focusses on whether the information is summarized effectively and no unnecessary information is included. This criterion focusses on audience awareness as a clear and efficient summary ensures that may result in time efficiencies for the recipient.
- **Genre and style:** this criterion is designed to focus the rater on the appropriateness of the tone and register to the purpose and audience of the document.

## Towards Improved Quality of Written Patient Records

- Organisation/layout/presentation: This criterion focusses on how well a document is organised and laid out, as well as the quality of the handwriting.
- Language: Our findings suggest health professionals are concerned with linguistic features only to the extent that they facilitate or obstruct retrieval of information. This criterion is designed to assess the quality of the language used in the document, and in particular focusses on the accuracy of the language and whether it interferes with reading comprehension.

As the data provided by the health professionals do not provide an indication of levels which could be suitable for the use of level descriptors in the rating scale, this first phase only prepared the highest level descriptors in the rating scale, drawing directly on the indicators developed in the previous phase. Table 5 below shows how this draft scale was laid out.

**Table 5.** Draft rating scale – top level descriptors

<b>Purpose</b>	<b>Content</b>	<b>Conciseness &amp; Clarity</b>	<b>Genre/style</b>
Purpose of document is clear throughout; document is effective in achieving purpose	Content is appropriate to intended reader and addresses what is needed to continue care (key information is included; no important details missing); content from case notes is accurately represented	Length of document is appropriate to case and reader (no irrelevant information included); information is summarized effectively and presented clearly	Writing is clinical/factual and appropriate to genre and reader (discipline & knowledge)

<b>Organisation</b>	<b>Layout/Presentation</b>	<b>Language</b>
Structure/organisation and paragraphing is appropriate, logical and clear to reader; key information is highlighted and sub-sections well organised	Document is legible and well laid out	Technical language, abbreviations and polite language are used appropriately for document and recipient; spelling/capitalisation, grammar and sentence structure are accurate and appropriate and allow a flow of language

### 7.2 Scale development workshop and pilot ratings - senior OET raters

We decided that to develop the remaining level descriptors, it was important to include experienced OET raters in the process. Two senior raters participated in this phase together with three project team members with expertise in language test development. The level descriptors were developed using a ‘bottom up’ process by carefully reviewing OET writing samples at different score levels, considering the qualities in these writing samples in relation to each of the criteria and formulating the descriptors. This involved a process of designing a descriptor, reading more OET writing samples (written in response to four different task types; two for nurses and two for doctors), revising the descriptors, considering the relationship of the descriptors to the adjacent levels and so on. The rationale for this approach was to ensure that the descriptors as closely resemble the discourse used by test takers as possible.

Following this workshop, the two OET senior raters and the three project team participants individually applied the rating scale to a further 20 OET writing samples from different score levels. Any differences in ratings were discussed following these pilot ratings and the agreed scores for these 20 scripts were stored for the larger rater training described below. Further minor adjustments to the rating scale were made at this point following comments by the raters. The final rating scale developed for the larger trial are found in Appendix 2.

### **7.3 Assessor training**

To help the OET language assessors understand and apply the new criteria, a rater training workshop was held at the OET Centre. Fifteen experienced OET raters agreed to participate in the training and the subsequent rating of 100 writing samples.

In advance of the workshop, the OET provided the project team with a large number of writing samples, written in response to four OET writing prompts (two for nursing and two for doctors).

At the workshop, assessors were briefed about the project and given an outline of how the level descriptors were created. Following this, the raters were provided with the rating scale (Appendix 2) and an accompanying document providing more information about each rating scale category (Appendix 3). Raters were also briefed in detail on each rating scale criterion. Raters were then given sample OET writing performances one by one, asked to rate the performances and then provided with the agreed scores. In total, the raters rated eight writing samples in the training session. For four of these the scores were provided prior to rating and for a further four the scores were only provided once the raters had applied the new rating criteria themselves. Following each script, there was a discussion in the group. We also provided the raters with a highlighted version of the case notes, where crucial information was highlighted in yellow and redundant information was highlighted in pink. This use of highlighting was done to ensure raters were consistent in identifying these pieces of information. However, we did not provide the raters with a sample answer to the four prompts.

Following the rater training workshop, raters completed a feedback questionnaire and were then provided with a pack of 100 writing samples which were selected from the writing samples provided by the OET Centre, drawing on scripts at all OET levels and written in response to four different tasks. Each rater was also provided with the prompts, the rating scale and description of criteria and a further questionnaire to be completed at home following the completion of the rating. The rating samples were distributed to raters so that all scripts were rated by at least three raters (in some cases four) in the group of fifteen participants. Each rater received a different sample of scripts, but we ensured that there was sufficient overlap between writing scripts to allow for appropriate and meaningful statistical analysis.

Following the completion of the rating of the 100 writing samples, each rater participated in a 30-minute interview with a researcher on the project team.

The rating data was subjected to a many-facet Rasch analysis using the statistical software FACETS (Linacre, 2016). With this analysis, we explored rater behaviour and the functioning of the rating scale as a whole, as well as the individual sub-scales.

To investigate whether the rating scale assessed separate abilities, we used principal axis factoring. Before this analysis, both the determinant of the R-matrix and the Kaiser-Meyer-Olkin measure of sample adequacy were calculated to ensure suitability of the data to the analysis. To determine the number of factors to be retained in the analysis, scree plots and eigenvalues were examined. Eigenvalues above 1 were retained.

The questionnaire data were summarized. The interview data (following transcription) as well as any qualitative comments on the questionnaires were subjected to a qualitative analysis, involving the identification of themes and sub-themes. This analysis involved a hermeneutic process of reading, analysing and re-reading of the data (Hycner, 1985). The coding themes that emerged during the process were then grouped into categories. The results of these analyses are described below.

### **7.4 Findings – OET raters application of professionally-relevant rating scale**

The functioning of the new, more professionally-relevant rating scale was established based on a statistical analysis of the ratings, as well as the analysis of raters' questionnaire response and thematic analysis of the post-rating interviews.

#### *7.4.1 Statistical analysis*

The statistical analysis set out to achieve several goals; (1) to examine whether raters were able to apply the scale consistently and without too much variation in terms of severity; (2) to examine rating scale functioning, both for the rating scale as a whole and at individual sub-scale level; and (3) to examine whether the scale tapped into an overarching construct or whether several sub-dimensions were identified, requiring discussions about score reporting.

##### *7.4.1.1 Rater functioning*

The analysis of rater functioning was conducted using many-facet Rasch analysis. The results showed that the raters differed in their severity from each other, but that this was no more than would be expected on any performance assessments (Eckes, 2011; McNamara, 1996). The most severe rater was rating about one score point harsher than the most lenient raters. The raters were all bunched within one logit either side of the mid-point of the logit scale.

In terms of consistency, we drew on the infit mean-square statistics from a many-facet Rasch analysis. High infit mean-square statistics provide an indication of raters rating erratically, more than predicted by the model. A rater is found to be rating inconsistently if their infit mean-square value is more than two standard deviations (SD) from the mean (Linacre, 2003). Only one rater was identified as misfitting using these criteria. Considering that the rater

training was very short and that the raters were applying the new criteria for the first time, the results were therefore encouraging.

#### 7.4.1.2 Rating scale functioning

We were also interested in how well the rating scale as a whole and the individual sub-scales functioned during the first large-scale trial. To examine this, we first examined the fit statistics of the six criteria. None of these were found to be misfitting. We also scrutinized the category statistics for the rating scale (Linacre, 2004; Bond & Fox, 2007). The requirement that the average measures advance monotonically with each higher band level was met by both the rating scale as a whole as well as all trait sub-scales. The Rasch-Andrich thresholds (the category thresholds) should also advance monotonically. This was generally also the case, although there were some instances where this did not hold at the very lowest level of some sub-scales, where very little data was available. The lowest scale category, Level 1, was rarely used by raters. This makes sense, as OET test takers are usually relatively proficient and rarely display features consistent with very low ability writers. We were also not provided any samples of scripts previously rated at Band E for this study. No mean-square outfit statistics for individual band scales were found to be higher than 2., showing that the data fit the model.

It was further of interest to us whether any of the sub-scales are able to discriminate between the test takers more than others. To arrive at the results of this analysis, we scrutinized the variance of the person ability for each criterion. Table 6 below presents the results of the analysis. It can be seen that ‘Purpose’ was the least discriminating sub-scale, while ‘Language’ was the most discriminating. It is interesting that the three criteria that were found to be the least discriminating related to the sub-scales assessing content, the area arguably least familiar to the raters, while the sub-scale that was found to be the most discriminating was ‘Language’, the trait with which most raters are most at ease.

**Table 6.** Sub-scale discrimination

<b>Criterion</b>	<b>Range (variance) of person ability</b>
Purpose	12.27
Content	13.74
Conc & Clarity	14.04
Genre & style	17.22
Organisation	17.48
Language	20.03

A further aspect we were interested in was whether the various sub-scales or traits were combining to measure one underlying construct or whether the analysis could identify multi-dimensionality in the data. We drew on three different analyses, to explore this question. Firstly, we examined a correlation matrix of the six criteria in our analysis. Table 7 shows that the correlations ranged from .417 to .624, indicating some relationship but little indication that any of the criteria are redundant.

**Table 7.** Sub-scale correlations

Criteria	PUR	CON	CC	GS	ORG	LAN
----------	-----	-----	----	----	-----	-----

## *Towards Improved Quality of Written Patient Records*

Purpose (PUR)	*					
Content (CON)	.599	*				
Conciseness and clarity (CC)	.568	.624	*			
Genre/style (GS)	.494	.499	.611	*		
Organisation/layout/presentation (ORG)	.484	.517	.625	.608	*	
Language (LAN)	.417	.431	.535	.584	.562	*

We then conducted an analysis using principal axis factoring to examine whether several components can be identified among the criteria. The analysis indicated that there was only one large factor with an eigenvalue of 3.839, accounting for 64% of the variance. No further large components were identified, indicating that the criteria were all working together to measure one underlying construct.

A final principal components analysis was conducted on the residuals of the criteria using Winsteps (Linacre, 2016). The eigenvalue of the first contrast (first PCA component in the residuals) was 1.76, which was only slightly higher than the value expected for random data (Raiche, 2005). Two sub-strands could be identified. Content and Purpose were the first sub-strand and Language and Genre/style formed the second. While these indicate that each are measuring slightly different things, these were not strong enough to be called separate dimensions (Linacre, March 2017, personal communication) and therefore do not justify separate score reporting.

### *7.4.2 Qualitative analysis*

The results from the post-training questionnaire as well as the post-rating questionnaire showed that only about half of the raters felt comfortable using the revised scale. This is not surprising considering that it was their first encounter with the revised descriptors and that the rater training was fairly short. The majority of the raters felt that the ‘rating scale helped them distinguish between test taker performances (60% of raters), that the scale reflects what health care professionals look for when reading written handover communication (87%) and that the layout of the scale was appropriate (87%). Their confidence in the ‘Language’ and ‘Purpose’ subscales was the highest (80% and 71% respectively) while only 53% were confident in their use of ‘Organisation/Layout/ Presentation’ and 40% in ‘Conciseness and clarity’.

The interviews showed that the raters varied in their assessment of the rating scale. Some thought that it was a considerable improvement on the current criteria while others were concerned about the shift away from language to medical communication, which some thought they were not qualified to assess. Raters suggested that they would need to be provided with more support if these descriptors would be implemented for operational rating, a suggestion which we will further elaborate on in the recommendations of this report.

Raters also commented on the individual sub-scales. Most raters liked the new criterion ‘Purpose’ as well as the inclusion of ‘legibility’ under ‘Organisation/Layout/ Presentation’. Two raters felt that the ‘Content’ and ‘Conciseness and clarity’ categories overlapped or were difficult to distinguish. The raters liked that they were provided with the highlighted tasks,

providing an indication of which aspects of the case notes should be included in the letter and which should be excluded. Some suggestions were made on improving this system which we discuss in the recommendation section. They also commented on the fact that they would have liked to draw on a sample answer while rating.

### **7.5 Summary of Phase Three**

The statistical results as well as the feedback from the raters indicate that raters can apply the more professionally-relevant rating criteria to the OET writing in a meaningful and consistent manner. The results showed that the criteria seemed to measure one underlying construct. In the interviews, the raters asked for more training and more guidance on certain aspects during the training to ensure that they are able to move from the focus on language to commenting on medical communication more broadly.

## **8. PHASE FOUR: SETTING MINIMUM STANDARDS FOR PROFESSIONAL REGISTRATION**

The final phase involved a standard setting exercise for the OET writing test using the judgements of clinical educators in medicine and nursing. The participants provided judgements on OET writing samples for which scores on the new rating scale were available as well as scores using the previous criteria. It was therefore possible to evaluate and compare the standard-setting judgements on the rating results from the two scales. The phase set out to answer the final research question: ‘What *minimum standards* should be set for professional registration of migrant health professionals?’

### **8.1 Methodology**

Several small workshops were conducted with both nurses and doctors as standard-setting subject-matter specialist informants. The data collected with the workshops with the medical professionals formed the data for Simon Davidson’s PhD thesis. He also collected additional data from doctors in the form of think-aloud protocols, which do not form part of the main funded study and are not therefore included in this report.

Eighteen doctors participated in the doctors’ workshops, drawn from a variety of sub-disciplines and contexts (GPs, specialists, consultants and medical educators). Some of the participants were also from non-English-speaking backgrounds. The mean years spent in their profession was 21. A number of the participants also had experience in supervisory positions, supervising junior, entry-level doctors. Eighteen nurses participated in the nursing standard-setting workshops. They were recruited from a range of contexts and sub-disciplines (e.g., intensive care, perioperative, community health, nursing education, general wards). On average, they had 17 years’ experience working in nursing. Almost all of them had experience in supervising new graduates and new entrants into the profession.

To set the minimum standards, we selected the analytic judgement method (Plake & Hambleton, 2011) for this project. The analytic judgement method has the advantage that participants are not presented with the writing samples in any particular order and do not know how the writing samples were previously rated. In this specific-purpose context, this has the advantage of not forcing the ranking of the writing samples on the participants, who may order the writing samples differently from the language-focussed OET raters. Workshop participants were asked to decide whether a writing sample they read is ‘strong’, ‘minimally competent’, ‘not yet competent’ or ‘unsatisfactory’ or whether it falls in any category between these broad groupings. The procedure and decision document was trialled on a small group of participants prior to the main data collection.

In each of the workshops participants were presented with a short introduction to the overall project. They were then asked to discuss orally in the group what it means for an overseas-trained health professional to be ‘minimally competent’ in written English communication skills in their respective workplace (this was already set as a homework task beforehand, so that the health professionals were prepared). Following this discussion, the participants were also asked to discuss what features in their writing would make a health professional ‘strong’, ‘not yet competent’ or ‘unsatisfactory’. Following this discussion, the participants were presented with six writing samples, which they each individually read and then placed into one of the main categories (strong, minimally competent, not yet competent, unsatisfactory) or one of the in-between categories. Once they had made their judgement, they discussed this in their groups. Any discrepancies were discussed and participants were able to change their ratings if they wanted to. There was no requirement to agree within the group, however. Following the workshop, each participant was given a take-home pack of a further 30 scripts to judge at home. Following the workshop as well as the in-house rating, the participants also completed a questionnaire. This was to elicit their reactions to the workshop and the task and their confidence in their judgements.

The quantitative workshop data were analysed using two methods. Firstly, we conducted a many-facet Rasch analysis using FACETS (Linacre, 2016) to ascertain how the standard-setting judges performed. In particular, we were interested in establishing whether any of the judges were rating differently to the group, in particular by being very harsh or lenient or inconsistent in their ratings. We expected differences in leniency and harshness between the judges because, other than in the training of judges rating language performances for large-scale tests, the workshop participants in this phase were not required to agree or rate like the other judges. The participants also all drew on experiences from different work contexts which may have slightly different expectations and requirements for written communication. We were, however, concerned about judges being inconsistent, that is, applying the rating categories differently to the other judges. Identifying any such judges was important to ensure that their judging behaviour did not adversely affect the standards set for the professions.

One judge from each profession was removed from each data set because of inconsistent rating behaviour. We tallied the OET writing scores provided by the OET raters (using the existing criteria) for any scripts that the remaining standard-setting judges placed in any in-between

categories. For example, any script that was placed in the category between ‘minimally competent’ and ‘not yet competent’ was tallied and the mean of the OET writing scores was calculated. The mean signified the new cut-score between ‘minimally competent’ and ‘not yet competent’. This process was repeated for all the other in-between categories to arrive at cut-scores between the different OET writing levels.

## **8.2 Findings**

The results section of Phase 4 aims to address a number of questions. First, do the standards set in the workshops (by the health professionals) differ from the current, operational standards? If so, would any test candidates be classified differently according to these new standards?

Second, do the standards set for nurses differ from those for doctors?

Third, would candidates, when scored against existing criteria, be classified into the same score categories if their performances were scored against the new criteria?

The findings are reported in three sections to directly answer these questions.

### *8.2.1. Findings – new standards compared to existing standards (medicine and nursing combined)*

Currently, the OET has one set of cut-scores which are applied across all professions equally. For this reason, we also calculated new, combined cut-scores, based on the data collected from both nursing and doctor workshops. Table 8 sets out the new cut-scores combined to the existing standards. It can be seen that the passing standard (between band B and C) is slightly higher when applying the results from our standard-setting workshops. The B band is much narrower and the A band much wider. Similarly, the C band is narrower, meaning it would be harder to get a C for those on the cusp between bands C and D.

**Table 8.** Combined medicine and nursing cut-scores (new vs existing)

<b>Medicine and Nursing combined</b>		
<b>Band</b>	<b>Score range new cut-scores</b>	<b>Score range existing cut-scores</b>
A	6.00 - 5.21	6.00 - 5.60
B	5.20 - 4.92	5.59 – 4.80
C	4.91 - 4.66	4.79 – 4.20
D	4.65 – 0.01	4.19 – 3.40
E	The range for E was not considered in this study	3.39 – 0.01

Table 9 shows the impact of these revised combined cut-scores on the distribution of band scores across the data set of 490 writing scripts that was used for our study. What can be seen in Table 9 is that the pass rate would be significantly lower if the new standards were applied using the existing rating criteria. While the pass rate using the existing cut-scores was 53.26% for this data set, if the new cut-scores were applied, the pass rate would only be 38.16%. Within the group that passes the OET, there would however be a higher percentage of test takers in the A band level and fewer in the B band.

**Table 9.** Combined medicine and nursing cut-scores – impact on score distribution

<b>Impact of combined medicine and nursing standards</b>				
<b>Band</b>	<b>Old rating scale old cut-scores</b>		<b>Old rating scale new cut-scores</b>	
A	65	13.26%	145	29.59%
B	196	40.00%	42	8.57%
C	157	32.04%	141	28.78%
D	70	14.29%	162	33.06%
E	2	0.01%	-	-

Some redistribution would also be seen in the band levels assigned to test takers not passing, with fewer receiving a C band and more being grouped into a D band level.

*8.2.2. Findings – new standards compared to existing standards (comparison of nursing and medicine standards)*

We conducted separate workshops for nurses and doctors to establish whether members of these two professions would set different standards. Table 10 sets out the new nursing standards (in comparison to the existing standards) and the same information can be seen in Table 11 for medicine. It can be seen that the passing standard (i.e., the cut-score between B and C) was set lower for the nurses (but still slightly higher than the current cut-score) when compared to that of the doctors (Table 11). The doctors’ passing standard set empirically in the workshops was substantially higher than it is currently on the OET.

**Table 10.** Nursing cut-scores (new vs. existing)

<b>Nursing</b>		
<b>Band</b>	<b>Score range new cut-scores</b>	<b>Score range existing cut-scores</b>
A	6.00 - 5.13	6.00 - 5.60
B	5.12 - 4.82	5.59 – 4.80
C	4.81 - 4.56	4.79 – 4.20
D	4.55 – 0.01	4.19 – 3.40
E	The range for E was not considered in this study	3.39 – 0.01

**Table 11.** Medicine cut-scores (new vs. existing)

<b>Medicine</b>		
<b>Band</b>	<b>Score range new cut-scores</b>	<b>Score range existing cut-scores</b>
A	6.00 – 5.25	6.00 - 5.60
B	5.24 – 5.03	5.59 – 4.80
C	5.02 – 4.77	4.79 – 4.20
D	4.76 - 0.01	4.19 – 3.40
E	The range for E was not considered in this study	3.39 – 0.01

Tables 12 and 13 show the impact on the pass and fail rates for the data set used in this study. In the case of the nursing data (Table X), 46.3% of nurses would have passed (i.e. received an A or B grade) using the existing cut-scores. With the new cut-scores, the pass rate would have been 38.49%. As was seen in the combined data set above, there would have been substantially more A grade passes than B grade passes and a higher proportion of D grades in the fail group.

**Table 12.** Nursing cut-scores – impact on score distribution

<b>Impact of Nursing standards</b>		
<b>Band</b>	<b>Old rating scale old cut-scores</b>	<b>Old rating scale new cut-scores</b>

## Towards Improved Quality of Written Patient Records

A	22	8.73%	62	24.60%
B	94	37.30%	35	13.89%
C	79	31.35%	66	26.19%
D	55	21.83%	89	35.32%
E	2	0.01%	-	-

The same trend as above can also be seen in the Medicine data set shown in Table 13. Using the existing descriptors and cut-scores, the pass rate on the data set used in this study was 60.93%. If the same descriptors were used but the new cut-scores would be implemented, the pass rate will be significantly lower, at just 43.69%. The same redistribution among test takers falling into A and B grades as well as C and D grades was also seen.

**Table 13.** Medicine cut scores – impact on score distribution

<b>Impact of Medicine standards</b>				
<b>Band</b>	<b>Old rating scale old cut-scores</b>		<b>Old rating scale new cut-scores</b>	
A	43	18.07%	81	34.03%
B	102	42.86%	23	9.66%
C	78	32.77%	52	21.85%
D	15	6.30%	82	34.45%
E	-	-	-	-

### 8.2.3. Findings – comparison of passing standards using new and existing criteria

**Table 14.** Cut-scores on new rating criteria compared to existing cut-scores – both professions

<b>Medicine and Nursing combined</b>		
<b>Band</b>	<b>Score range new cut-scores on new rating criteria</b>	<b>Score range existing cut-scores</b>
A	6.00 - 4.67	6.00 - 5.60
B	4.66 - 4.37	5.59 – 4.80
C	4.37 - 4.20	4.79 – 4.20
D	4.19 – 0.01	4.19 – 3.40
E	The range for E was not considered in this study	3.39 – 0.01

Table 14 above sets out the corresponding cut-scores between the OET grade levels for the new rating scale. These are quite different and indicate how differently the raters engaged with the new rating criteria. Tables 15 and 16 below set out the same cut-scores for medicine and nursing separately. The lower cut-scores were reflected in the rater interviews in Phase 3, where raters reported feeling liberated from the existing criteria and the sense that they should not award scores below 4. The scores awarded in Phase 3, although on the same number of band levels, were substantially lower.

**Table 15.** Cut-scores on new rating criteria compared to existing cut-scores – medicine only

<b>Medicine</b>		
<b>Band</b>	<b>Score range new cut-scores on new rating criteria</b>	<b>Score range existing cut-scores</b>
A	6.00 - 4.65	6.00 - 5.60
B	4.64 - 4.43	5.59 – 4.80
C	4.42 - 4.25	4.79 – 4.20
D	4.24 – 0.01	4.19 – 3.40
E	The range for E was not considered in this study	3.39 – 0.01

**Table 16.** Cut-scores on new rating criteria compared to existing cut-scores – nursing only

<b>Nursing</b>		
<b>Band</b>	<b>Score range new cut-scores on new rating criteria</b>	<b>Score range existing cut-scores</b>
A	6.00 - 4.68	6.00 - 5.60
B	4.67 - 4.35	5.59 – 4.80
C	4.34 - 4.15	4.79 – 4.20
D	4.14 – 0.01	4.19 – 3.40
E	The range for E was not considered in this study	3.39 – 0.01

### **8.3 Summary of Phase Four**

Overall, this phase was successful in establishing new cut-scores, which were arrived at empirically drawing on the judgements of subject-matter experts from two domains, medicine and nursing. If the cut-scores were to be adopted, the pass rates for both professions would be reduced, with some additional shifts in the band allocation for both the passing and the failing group. It was also interesting to note that the standards set on the medicine sub-tests were more stringent than those on the nursing test. This requires a discussion of whether different professions should have different passing standards. If the new cut-scores were adopted, it might result in fewer concerns by health professionals that overseas-trained health professionals are not yet ready to practice in Australian health care settings. We also calculated the corresponding cut-scores if the new rating scale was adopted to ensure this data is available if Recommendation 1 of this report is implemented.

## **9. RECOMMENDATIONS**

### **9.1 Recommendations for Practice**

There are a number of recommendations from this study which relate to the delivery of the OET writing sub-test.

*Recommendation 1: We propose that the new professionally-relevant writing criteria are adopted for the OET writing sub-test*

This recommendation is based on the outcomes of the current study, which elicited health professionals' values of written handover communication and converted these into rating criteria which can be used by the language-trained OET raters in judging performances on the OET writing test. The results showed that despite a short training session, the raters were able to reliably apply the new rating criteria to OET performances. The new rating scale expands the construct of written communication as measured by the OET writing task and is therefore an important advance in the assessment of written English for health care professionals. While raters need more training and support to implement the criteria (see recommendations below), the trial results were promising.

*Recommendation 2: We propose the OET revise the existing specifications for the OET writing task to include scenarios and contextual information needed to elicit the qualities of communication valued by participants but currently not included in the writing task.*

A careful scrutiny of the descriptors written to represent the values of the health professionals revealed that not all could be incorporated into the current OET criteria due to limitations in the current OET writing task specifications. Most of these descriptors could be incorporated into an elaborated set of test specifications of the writing test, with the effect of broadening the abilities assessed by the task and most likely yielding positive washback on test takers' understanding of what is required in the domain and on test preparation activities. We recommend that the following aspects are included into the test specifications:

### 1. Information about the context of the writer and the recipient

Each writing task could include information about both the writer and the recipient of the document, including what each already knows about the patient, the professional background, speciality and possible hierarchy between the writer and the recipient and (if applicable) whether and how the patient has been involved in any decision-making regarding their care.

### 2. Inclusion of problematic point in task specifications

We further recommend that the test specifications have provisions for including a problematic point in the writing task or case notes. This should enable test takers to demonstrate the ability to express doubts or uncertainty about a patient's symptoms or diagnosis, a feature mentioned by the health professional workshop participant when reviewing the written handover letters.

### 3. Audience and purpose of writing

We recommend that the OET consider extending the writing specifications to include the possibility that writing tasks are written for multiple audiences and for multiple purposes. This was found to be commonly the case in the real-world healthcare domain and would enable writers/ test takers to demonstrate that they can, for example, write at the same time to a professional colleague as well as adding information that is important for a patient or their family members, or for health information services staff.

### 4. Language requirements

It was clear from the sample letters collected from the patient records in Phase 1 as well as the comments on the qualities from health care professionals in Phase 2, that it is acceptable in the domain to use bullet-points and sentence fragments in the types of documents we examined. While it is clear that sustained use of these features would make the assessment of the OET writing task difficult, we recommend that the test specifications allow for use of bullet-points at certain points in the documents (e.g., when previous conditions or medications are listed, or when the requirements for follow-up are described). Allowing this flexibility would have no direct implications for the rating criteria we developed (as these contain no stipulation that complete sentences be used), but the option of using bullet points where

appropriate would need to be included in the training of OET raters as well as in the instructions to candidates.

*Recommendation 3: We recommend that **domain experts** be involved in the **training of OET writing raters** and available during operational rating*

Feedback from raters involved in trialling the new rating criteria showed that raters require additional support to implement more professionally-relevant criteria. We therefore recommend that one or more domain experts be available when raters are trained to answer domain-specific questions. For example, during the trial it became clear that raters were not always sure about the expected discourse structure of certain response types and we therefore recommend including more guidelines on this area in the training. Domain experts could discuss acceptable organisational structures (such as the ISBAR or SOAP<sup>2</sup> protocols) with the raters during training to raise their awareness of accepted conventions in health care contexts. Similarly, there were many questions about which specific information in the case notes should be included into the response and which aspects should be left out, as they were redundant. While the trial of the highlighted task materials (see Recommendation 4) was helpful, this did not completely alleviate all concerns and it would therefore be helpful to have domain experts available to help with any such content-related questions. We further feel that having domain experts available or ‘on call’ to answer crucial questions about specific tasks during rating periods would be advisable.

*Recommendation 4: We recommend that additional **support materials** be offered to raters during operational rating periods*

After a successful trial in Phase 3 of this study, we recommend that the OET adopt the practice of providing raters with information about key and unnecessary information in the case notes to support their rating activities. Each of these areas was highlighted in a different colour during the trial and the raters suggested during the interviews that absolutely necessary information could further be presented in bold font. The raters were generally very positive about this practice. The list of indicators provided in Table 3 may also be useful for test developers when preparing the tasks and the support materials for raters.

We also recommend the continuation of the current practice of providing raters with a sample response. It is important however that the content included in this response matches with the highlighting in the case notes (see above). Furthermore, it is important that raters understand that the discourse structure in the sample response is not necessarily the only possible way the response could be organised. Again, we refer to Recommendation 3, the inclusion of domain expert input during training, to support any discussion about what kind of variations from the sample response might be acceptable.

---

<sup>2</sup> ISBAR: Identify, Situation, Background, Assessment and Recommendation  
SOAP: Subjective data, Objective Data, Assessment, Plan

*Recommendation 5: We recommend that the OET inform test candidates and test preparation centres of the new rating criteria and make available training materials which reflect the qualities embodied by these criteria*

We also recommend that the rating criteria are published on the OET website and included in any training materials to ensure all stakeholders understand the criteria used to judge the writing performances.

To ensure that the OET writing task results in positive washback on the teaching and learning which takes in preparation for the test, we recommend that training materials are published which are easily accessible to both test candidates and test preparation providers. In particular, we recommend that the materials include detailed guidance on the qualities of effective written communication in a richer context than possible on the test itself so that the importance of patient-centredness as a core value of health professions can be signalled and positive washback is achieved. Furthermore, such materials should clearly signal to potential test takers what criteria will be used to assess their performances and how these link with expectations of performance in the real- world domain.

*Recommendation 6: We recommend that prior to implementation of any changes to the OET rating criteria, the test specifications or the test tasks, key professional groups be involved in a consultation process.*

This research only focussed on two professional groups which directly contribute to, or access patient records, namely, doctors and nurses. While we think the proposed changes to the criteria are equally applicable to the other professions covered by the OET, we recommend that their suitability to the broader healthcare context is explored in consultation with key professional groups covering those professions (see also Recommendation 9 for further research) as well as with members from professional boards representing each profession.

*Recommendation 7: We recommend that the new cut-scores are implemented for nursing and medicine*

The results of the standard-setting workshops made it clear that, according to our subject matter expert participants, the current pass mark is too lenient. This also mirrors some indications from the professions that overseas-trained health care professionals are entering Australian workplaces without sufficient language skills. We therefore recommend that the new cut-scores are implemented for nursing and medicine.

We further recommend that the different nursing and medicine cut-scores are adopted, making it slightly easier for nurses to pass the OET. This is a reflection of the differing workplace demands of the two professions.

## **9.2 Recommendations for Research**

Apart from the recommendations for practice we discussed above, we would also like to make a number of recommendations for further research arising from this project.

## ***Towards Improved Quality of Written Patient Records***

*Recommendation 8: We recommend that the new rating criteria are **verified and trialled across a wider range of tasks** for medicine and nursing and across tasks from all other ten professions*

For the purposes of this study, OET raters rated OET writing performances based on four task scenarios (two from nursing and two from medicine). We recommend that the new rating criteria are verified and trialled across more tasks, including tasks from other professions to ensure that the criteria apply more widely.

We also recommend canvassing the opinions of health professionals from the ten professions not included in this study as to the applicability of the revised descriptors to writing relevant to their discipline.

*Recommendation 9: We recommend that **additional standard-setting panels** are convened to set cut-scores for the **other professions***

The current study only focussed on two professions: nursing and medicine. To ensure the cut-scores implemented are representative of all twelve professions of the OET, we recommend that further standard-setting panels are convened and the impact of these new cut-scores on the pass rate is considered.

## 10. REFERENCES

- Barton D., Hawthorne L., Singh B., Little J. (2003). Victoria's dependence on overseas trained doctors in psychiatry. *People and Place* 11(1):54-64.
- Bond, T. and C. Fox (2007). *Applying the Rasch model. Fundamental measurement in the Human Sciences*. New York, Routledge.
- Dorgan K., Lang F., Floyd M., Kemp E. (2009). International Medical Graduate–Patient communication: A qualitative analysis of perceived barriers. *Acad Med*. 84(11):1567-75.
- Eckes, T. (2011). *Introduction to Many-Facet Rasch Measurement*. Frankfurt, Peter Lang.
- Elder, C., McNamara, T., Woodward-Kron, R., Manias E., McColl, G., Webb, G. (2013). Towards improved healthcare communication: Development and validation of language proficiency standards for non-native English speaking health professionals. Final report for the OET Centre. Language Testing Research Centre, University of Melbourne.
- Woodward-Kron R, Fraser C, Pill J, Flynn E. (2014). How we developed *Doctors Speak Up*: An evidence-based language and communication skills open access resource for International Medical Graduates. *Medical Teacher*, 37(3): 31-33.
- Garling P. (2008). Acute Care Services in NSW Public Hospitals. Sydney: NSW Government
- Konno R. (2006) Support for overseas qualified nurses in adjusting to Australian nursing practice: a systematic review. *International Journal of Evidence Based Healthcare*, 4:83-100.
- Hawthorne L. (2012). International medical migration: what is the future for Australia? *Medical Journal of Australia Open*. 1 Suppl 3:18-21.
- Jacoby, S. (1998). Science as performance: Socializing scientific discourse through conference talk rehearsals. Los Angeles, Unpublished doctoral dissertation, University of California.
- Jacoby, S. & McNamara, T. (1999). Locating competence. *English for Specific Purposes* 18 (3), 213–241.
- Linacre, J. (2003). "Size vs. significance: Infit and outfit mean-square and standardized chi-square fit statistics." *Rasch Measurement Transactions* 17: 918.
- Linacre, J. M. (2004). Optimizing rating scale effectiveness. *Introduction to Rasch measurement*. E. V. Smith and R. M. Smith. Maple Grove, MN, JAM Press: 257-278.
- Linacre, J. M. (2016). Facets Rasch measurement computer program. Chicago, Winsteps.com.
- Linacre, J. M. (2016). Winsteps Rasch measurement computer program. Beaverton, Oregon, Winsteps.com.
- Manderson B., McMurray J., Piraino E. (2012). Navigation roles support chronically ill older adults through healthcare transitions: a systematic review of the literature. *Health Soc Care Community*, 113-127.

## ***Towards Improved Quality of Written Patient Records***

- Manias E., Jorm C., White S. (2008). Handover. How is patient care transferred safely? In: Jorm C. (Ed.). *Windows into Safety and Quality in Health Care*. Sydney: Australian Commission on Safety & Quality in Healthcare, p. 37-48.
- McNamara T. (1996). Measuring second language performance. London: Longman;
- Mullan F. (2005). The metrics of the physician brain drain. *N Engl J Med.* (17):1810
- Plake, B. and R. Hambleton (2001). The analytic judgement method for setting standards on complex performance assessments. *Setting performance standards: Concepts, methods, and perspectives*. G. J. Cizek. Mahwah, NJ, Lawrence Erlbaum: 283-312.
- Raiche, G. (2005). Critical eigenvalue sizes (variances) in standardized residual principal components analysis (PCA). *Rasch Measurement Transactions* 19(1): 1012.
- Standing Committee on Health and Ageing (2012). Lost in the Labyrinth: Report on the inquiry into registration processes and support for overseas trained doctors. Canberra: Commonwealth of Australia.
- Woodward-Kron R., Stevens M., Flynn E., (2011). The medical educator, the discourse analyst, and the phonetician: A collaborative feedback methodology for clinical communication. *Acad Med.* 85(5):565-70.



## **APPENDIX 1: RATING SCALE USED FOR LARGE-SCALE TRIAL WITH OET RATERS (PHASE 3)**

The scale has been removed for test security reasons



## APPENDIX 2: DESCRIPTIONS OF SCALE CRITERIA

Criterion	Description
<p><b>Purpose</b></p> <ul style="list-style-type: none"> <li>- Helps the reader get a <b>quick and precise sense</b> of what is asked of them</li> </ul>	<p>Due to time constraints, health care professionals want to understand the purpose behind a written handover document (e.g. referral letter) very quickly and efficiently. This criterion therefore examines how clearly the writing communicates the purpose of the document to the reader. The purpose for writing should be introduced early in the document and then clearly expanded on later (often near the end of the document). The purpose should be highlighted to the reader, so there is no need to search for it.</p> <p>For example, a writer might at the beginning of the letter write ‘I’m writing to you today to refer patient X who is now being discharged from hospital into your care’. Later in the letter, specific instructions for the health care professional on continuing care should be listed.</p>
<p><b>Content</b></p> <ul style="list-style-type: none"> <li>- Considers <b>necessary information</b> (audience awareness: what does the reader need to know?)</li> <li>- Considers <b>accuracy</b> of information</li> </ul>	<p>The content criterion examines a number of aspects of the content:</p> <ul style="list-style-type: none"> <li>• All key information is included</li> <li>• No important information is missing</li> <li>• Information is accurately represented</li> </ul> <p>Audience awareness is key here. The writing needs to be appropriate to the reader (and their knowledge of the case) and what they need to know to continue care;  <i>Please refer to the accompanying documents for a list of the key information that should be included for each task</i></p>
<p><b>Conciseness &amp; Clarity</b></p> <ul style="list-style-type: none"> <li>- Considers <b>irrelevant information</b> (audience awareness: what doesn’t the reader need to know?)</li> <li>- Considers how <b>effectively</b> case is <b>summarized</b> (audience awareness: no time is wasted)</li> </ul>	<p>Health care professionals value concise and clear communication. This criterion, therefore, examines whether unnecessary information from the case notes is included and how distracting this may be to the reader (i.e. does this affect clarity). Is there any information that could be <u>left out</u>? It also assesses how well the information (the case) is summarized and how clearly this summary is presented to the reader.</p>
<p><b>Genre/style</b></p> <ul style="list-style-type: none"> <li>- Considers the <b>appropriateness</b> of features such as <b>register and tone</b> to the document’s purpose and audience</li> </ul>	<p>Referral letters and similar written handover documents need to show awareness of genre by being written in a clinical/factual manner (leaving out, e.g. personal feelings and judgements) and awareness of the target reader through using professional register and tone. The use of abbreviations should not be overdone and assume common prior knowledge - if written to a medical colleague in a similar discipline, then abbreviations and technical terms would be entirely appropriate, but if the medical colleague was in a totally different discipline, or a letter was from a specialist to a GP, more explanation and less shorthand would be desirable. As well, if the target readership could also include the patient, the information must be worded appropriately to the patient, e.g. medical jargon would be inappropriate.</p>
<p><b>Organisation/Layout/Presentation</b></p> <ul style="list-style-type: none"> <li>- Considers <b>organisational features</b> of the document</li> <li>- Considers <b>handwriting</b></li> </ul>	<p>Health professionals value documents that are clearly structured so it is easy for them to efficiently retrieve relevant information. This criterion examines how well the document is organised and presented. It examines whether the paragraphing is appropriate to the genre, whether sub-sections within the document are logically organised and whether key information is clearly highlighted to the reader so that this is not easily missed. The criterion also considers whether the layout of the document is appropriate and the handwriting legible.</p>
<p><b>Language</b></p> <ul style="list-style-type: none"> <li>- Considers aspects of language proficiency such as <b>vocabulary</b> and <b>grammar</b></li> </ul>	<p>Health professionals are concerned with linguistic features only to the extent that they facilitate or obstruct retrieval of information This criterion examines whether the language used is accurate and does not interfere with reading comprehension or speed.</p> <p><b>Please note:</b> unlike the current OET rating scale, this criterion does not consider the complexity of the language used as complexity was not mentioned as something valued by our health professional participants.</p>