BOOK REVIEWS

ILHA DO DESTERRRO

Models of working memory: Mechanisms of active maintenance and executive control (1999) . Miyake, Akira. Shah, Priti (Eds.). UK: CUP, xx + 506

By Augusto Buchweitz

'What is working memory, anyway?' Despite the seminal definition (Baddeley & Hitch 1974; Daneman & Carpenter 1980) of working memory (WM) being the system or mechanism at the basis of maintaining task-relevant information during the performance of a cognitive task (p. 1), and despite the familiarity of the term, addressing the question of what WM *really* is does not seem to be in the same ballpark of easily-put answers; it is not even in the same league. With the widespread interest on WM research, the literature at times has presented contradictory claims with diverse perspectives on nature, structure, and functions of WM. Thus, the above seemingly embarrassing, and yet fundamental, question lies at the basis of organizing this book.

In the interest of defeating misinterpretations and of developing some unification, eight sets of commonly-motivated issues were proposed to an all-star team of contributors. The contributors later met in a symposium entitled "Models of Working Memory" (1997) to present and exchange their ideas. What resulted was a scholarly, clarifying discussion on WM. The contributions may thus serve as a guiding star in a constellation of previous works. But unification was not the only concern of the editors and contributors: Graduate students and researchers alike were asked for reviews on different chapters (p. xvi). Thus clearness, or even simplicity, also played a hand in the making of the book.

In the following paragraphs, I will try and review some interesting aspects of the book; I will not tackle all eight questions proposed, considering, not only, space constraints, but also that the contributors themselves summarize the answers more suitably than I ever would.

Chapter 1

The introductory chapter presents the theoretical or modeling perspectives that motivated each of the eight sets of questions: (1) Basic mechanisms and representations in WM; (2) The control and regulation of WM; (3) The unitary vs. non-unitary nature of WM; (4) The nature of WM limitations; (5) The role of WM in complex cognitive activities; (6) The relationship of WM to long-term memory (LTM) and knowledge; (7) The relationship of WM to attention and consciousness; (8) The biological implementation of WM.

Chapter 2

Baddeley & Logie, following an initial presentation of the multiple-component model, provide overlapping answers to the theoretical questions. It is striking that the authors further developed the seminal 1974 model by Baddeley & Hitch, possibly trying to accommodate the hallmark central and subsidiary systems to more recent findings. One of the major changes is that the executive no longer holds a supplementary storage capacity beyond that of the slave systems (in the original model, it represented a pool of resources that could be used for control processes and storage purposes)–an alteration both theoretically (central executive too powerful) and empirically (WM can utilize temporary storage in systems other than the slaves) motivated (p. 37-8). Further changes to the specialized peripheral processors (loops) and the executive point to fractionation of the components. Moreover, Baddeley & Logie accept that simply resorting to a central structure without fully specifying the organization of central executive processes is not satisfactory (p. 39); clearly, there is room for future studies (p. 40). The chapter also provides a remarkable discussion on WM and complex cognitive activities, which others (for example, chapter 11) recognize not including as of yet.

Chapter 3

In chapter 3, Nelson Cowan presents his embedded-processes model of WM, in which great importance is assigned to mechanisms of activation and attention. One of the central features of the model postulates three hierarchically-arranged faculties at the source of WM information (the model clearly takes after its name): The larger component, LTM, encompasses the second, smaller component, activated memory, that encompasses the yet smaller component, the focus of attention (it is not possible to attend to information that is not activated!) (p. 63). In the illustration provided (p. 64) the focus of attention resembles an egg yolk, the inner core central to WM. The activated memory, in turn, resembles the egg white, just outside the core.

Whereas the model shapes a broader-to-more specific hierarchy of the components, the pieces of information more readily accessible to WM, inversely, are in the smaller structure. The modeled distance between the focus of attention and LTM illustrates that more effort is required to retrieve information from LTM (the pieces of the eggshell?). There is a clear similarity between the present model and that in chapter 2: The central executive. The difference, however, is that instead of controlling and regulating subsidiary systems, here the executive controls the focus of attention on different channels, setting attentional limits to WM based on time and capacity constraints. The time constraint, for example, indicates that the focus of attention decays—one cannot attend to a specific item forever. Capacity constraints indicate the need to allocate

WM resources more efficiently and profitably. The issue of attention (which is part of the common theoretical questions) is, thus, central to this model.

Chapter 4

Engle, Kane, & Tuholski model WM is based on controlled attention, general fluid intelligence (*gF*), and functions of the prefrontal cortex (PFC). WM capacity, in this chapter, translates STM plus controlled attention capacity (p. 104). The authors clearly blaze the trail to be followed in the discussion: Individual differences in WM actually mirror one's capability to control processing and attention (p. 104). To the point, much like chapter 3, focusing and controlling attention is at the basis of more optimal activation and maintenance of representations in WM.

WM, STM, and *gF* measures were studied to relate WM and controlled attention. Roughly, the researchers aimed at assessing whether the differentiation between WM and STM is necessary, and whether WM (after the common variance to WM and STM is partialled out) would correlate with *gF* in that both resort to the same component while in higher-order functioning (or complex tasks)—that is, controlled attention (p. 108). In short, the findings indicated that WM and STM constructs are highly related, but separable; and that the relationship between WM and *gF* (controlled attention) is highly significant (p. 110).

An extensive discussion on the biological implementation of the two constructs and that of controlled attention is provided. The authors anticipate that the experiments will show that PFC is associated with WM functions in tasks that require storage and processing of information, and with controlled attention tasks that require maintenance and focusing of attention (p. 117). Though seemingly fuelling the initial question of what is WM anyway, the three theorists based their argument on sound biological evidence. The only way out of this (apparent) predicament is, thus, to keep reading the book.

Chapter 5

Lovett, Reder, & Lebiere offer a goal-driven and goal-limited computational model of WM (ACT-R). The objective of computational modeling is to try and specify WM mechanisms in, and fit these mechanisms to programmable ways; to put it more simply, make them mechanically explicit (p. 136). The model establishes clear activation limits to WM: The latency for retrieving information is a function of the activation power of time (or, the *activation*th power of time); in other words, in keeping with a computational model—in computers, the latency is the time required to locate the first bit or character in a storage location—the lower the activation, the greater the latency (p. 142).

WM capacity in this architecture is limited by source activation and decay of activation, thus at the underlying architecture (p. 146) (mathematical equation on page 145). Metaphorically, WM capacity in this model resembles a

source of power and the goal nodes, lightbulbs connected in series. Thus, in more complex tasks, the power is spread more thinly over the 'bulbs': The light on goal nodes dims.

Page 155 offers a comparison between performance predicted by the model and that of participants, according to task complexity. The results show that as the complexity of an algebra problem increases, so does the predicted and actual time to solve the problem at the group level. In trying to address individual differences in WM, a defined parameter in the model (*W*, or attentional resource) was randomly varied based on a normal distribution (rather than based on different participants' data). Consequently, random variability in the equation seemed to better fit individual differences in performance (p. 164). At length, the mind-boggling and apparently invaluable aspect of this chapter, in my opinion, is that a fixed, programmable model variably captures human performance at the individual and group levels.

Chapter 6

Contributors Kieras, Meyer, Mueller, & Seymour put forth a model that applies a simulation software system: EPIC (Executive-Process/Interactive Control). The software runs a replication of human performers in different tasks. According to the writers, improving comprehensive modeling of human performance in complex, time-pressed tasks and environments (e.g. fighter plane cockpits) goes hand-in-hand with further understanding WM.

In humanoid fashion, EPIC operates perceptual processors of visual, auditory, and tactile input, and, also, motor processors for moving the hands, eyes, and speech articulators (a simulation software is available on the Internet; see also chapters 7 and 9 for online information). The auditory perceptual processor, for example, can encode words and strings of words that are subsequently entered in an artificial WM to model human performance in WM tasks. In this sense, EPIC-architecture WM stores and processors can be used to model a phonological loop (p. 191) with processing of stimuli and subvocalization (covert output), and rehearsal of the stimuli. The EPIC produces correct and incorrect recalls. The question is, thus, whether patterns of incorrect recalls match those of humans. The answer: Yes. What sticks in this chapter is that performance predicted by EPIC did equate with that of studies with humans. In future research, however, the contributors indicate the need to apply more complex tasks to the model (p. 218).

Chapter 7

In this chapter, Young & Lewis tackle the eight common questions from the perspective of another production system architecture: *Soar*. Chapters 5, 6, and the present chapter model WM alike, that is, by applying architectures to the WM psychological construct. The hallmark of the Soar architecture is the emphasis on WM and LTM involvement (long-term WM).

Three important components of the model are (1) 'Soar's dynamic memory' (SDM) or Soar's WM, which holds and controls information pertaining to the task in hand; (2) the 'production memory,' the LTM of the architecture (holds persistent knowledge); and (3) the *learning* mechanism, in which learning occurs while solving stalemates, or problems in processing (called *impasses*). New production rules learned (during a task) are stored in production memory (LTM) and, thus, may be later retrieved into the architecture's WM, or SDM to fire a rule or mechanism that avoids the previously-solved impasse (p. 230-1)–Soar operates a constant interaction between a fraction of production memory and SDM.

Three particularly interesting studies are discussed in the chapter. In one (Rieman, Young, & Howes, 1996), a Soar model named IDXL is applied in an exploratory, problem-solving task (browsing an unfamiliar computer application). The model learns by browsing different items (for example, menu icons) in the application; what is learned may be later retrieved from LTM—during exploratory problem solving—to SDM. In brief, IDXL models how humans can overcome WM constraints retrieving previously learned information from a larger-capacity component, LTM, into a capacity-constrained WM.

Chapter 8

Contributors Ericsson & Delaney also bring to the fore the notion of long-term WM. According to the two theorists, WM is not limited by a fixed capacity; rather, it is limited by skill and experience and, thus, the ability to efficiently and selectively encode, when needed, information stored in LTM (p. 257). Ericsson & Delaney underscore that the difficulty in finding a common ground in WM theory may be traced back to more fundamental constructs, such as those of LTM and STM. It seems thus that controversies in replying to the original question 'What is WM anyway?' are, actually, to be expected (p. 258) and that theoretical differences are possibly not being weeded out at the root.

But not all is lost. Ericsson & Delaney put forth an illustrative and elucidating discussion of three types of approach to WM research: (1) basic-capacity; (2) transient-storage; and (3) expert-performance. The discussion, in turn, seems to merge with their advocating skill and experience at the basis of individual differences in WM capacity. Studies related to (2), for example, show that experts' performing memory tasks within their 'realm' (e.g. chess masters recalling chessboard-piece positions after brief exposure) are apparently minimally waylaid by interjecting, unrelated tasks (p. 266)—information later retrieved into WM was stored in LTM (LT-WM)! One of the central issues of the chapter is that successful performance in skilled everyday activities that involve comprehension and semantic encodings little has to do with STM (or ST-WM) capacity; rather, successful performance is apparently pegged to efficiently and appropriately

retrieving information from LTM into WM (p. 285).

Chapter 9

In this chapter, Barnard armors his discussion with the framework of Interacting Cognitive Subsystems (ICS). ICS is a multiprocessor architecture pillared by three parallel types of processing: Direct processing; buffered processing—representations are 'burned' onto the image record that, in turn, allows for online use of the information by other subsystems (any similarity to WM is NOT a mere coincidence); and reuse, revival of stored information (p. 305-7). The processes, in turn, may run in nine different subsystems: Five peripheral, four central. The earlier are employed in representing information, the latter, transforming it (e.g. breaking words into phonemes). The latter two thus differentiate use of more recent and past input. Where does, in this sense, WM fit into this picture? Or, better yet, system.

First of all, ICS models a flow of information and representations between the subsystems once input is received, for example: In a WM span task, there may be a flow of input from peripheral (sensory) to central subsystems and back to peripheral subsystems (effector, for overt recall) (p. 320). A basic constraint to such flow across modes is that more than one subsystem (e.g. in performing dual tasks) may require drawing on the buffered mode of processing; thus, the architecture will adaptively reconfigure in ways that may or not be optimal (WM limitations arise from the interactions in the system!) (p. 324). Consequently, that the ICS does not postulate a central executive indicates that the type of input or output may determine resource allocation. Also, that the ICS postulates subsystems running parallel processing apparently accommodates findings of different memory constructs functioning separately (intact LTM in patients with impaired STM) (p. 332). At length, though the ICS was not originally devised to study WM, it does accommodate the issues proposed.

Chapter 10

Schneider presents a computational model, named CAP2, which, similarly to the ICS, for example, was not initially intended for modeling WM. Rather, CAP2 was intended to model human performance and skill acquisition. The architecture is based on layers of hundreds of thousands processing modules (p. 343). Whereas others set foot on theoretical constraints that are later applied, or accommodated to biological data, the CAP2 apparently runs in the opposite direction. In this sense, throughout the chapter Schneider models 'biological candidates' for different CAP2 components and functions.

Mapping the discussion back onto the issues of the nature of WM, CAP2 theorizes a central executive that captains the processing modules based on compressed information; in other words, for optimal reasons the executive

exerts its control based on 'reports' from the modules (p. 349). The control and regulation functions of the executive (p. 349-56) comprehend actions and activities central to the discussion on WM; for example: Buffering information; building new associations; monitoring; among others. Schneider places emphasis on the similarities and differences between the models in the book, especially the computational models Soar, ACT-R, and EPIC. However, at surface level the more striking similarity is the modeling of a central executive and peripheral processors (loops) in ways resembling those of Baddeley & Logie's multiple-component model (chapter 2) (p. 359). Schneider at last remarks being struck by the convergence of the models (p. 372)—thus in line with the author's comparative effort.

Chapter 11

This chapter presents O'Reilly, Braver, & Cohen's biologically based computational model of WM. As in other chapters, the contributors provide a description of the model as a stepping stone to the discussion of the theoretical questions. Three functional dimensions spearhead the model, together with a primary brain system counterpart: (1) Rapid learning of arbitrary information, of novel associations (hippocampus, or HCMP); (2) dynamic and robust active memory, maintenance of representations (prefrontal cortex, or PFC); and (3) sensory and motor processing (posterior and motor cortex, or PMC) (p. 382-3). The contributors do not deny that there are several other brain systems to be studied; nor that there are other functions. But the referred brain systems and their functions are central to the discussion on the nature of WM.

In this sense, for example, the PFC is at the heart of maintenance mechanisms in WM: It has the ability to maintain activation over time and to update representations (p. 399). However, the brain mechanisms and processes are a result of PFC, HCMP, and PMC interaction—thus the PFC is NOT to be mistaken for a central executive. Likewise, WM limitations are a result of a trade-off between the interacting but also competing brain systems (p. 400). The model fosters a neurobiologically plausible discussion of WM and draws on a complex discussion that ranges from anatomy to neuropsychology (p. 404).

Chapters 12 and 13

The two final chapters provide the ingredients to try and bring the discussion to a close. Chapter 12, in this sense, employs Kintsch, Healy, Hegarty, Pennington, & Salthouse in critically reading the answers provided. The chapter successfully integrates the answers provided over the book; there is, clearly, some compromise. For example, in relation to the control and regulation of WM, that some models do not postulate a central executive—but rather control emerging from the dynamic interaction between processes, goals, and impasses—does not leave

control out of the equation. No one thus really avoided resorting to some type of control.

Finally, chapter 13 addresses emerging consensuses, unresolved issues, and future research directions. It also declares the demise of older ideas, such as WM being modeled as a box, or a place in the brain (p. 448-9); rather, WM represents brain areas WORKING together to produce different phenomena. Miyake & Shah thus step in as the 'closers' of the original challenge. The discussion on unresolved issues and future directions helps fuelling a desire for more information. Thus, I will not spoil the fun of reading this last chapter; rather, I would like to conclude remarking that the book hits a grand-slam in, one, fostering and, two, documenting progress in the field, three, pitting theorists against each other not to find one better, but to find a gray area, and four, in setting a sound foundation for the study of WM. To the extent of my limitations, I would recommend reading the book to graduate students, and researchers alike, whose interests involve not only WM, but also other fields of interest in cognition.

References

Baddeley, A. D., & Hitch, G. J. (1974). Working memory: In G.H. Bower (Ed.) , *The psychology of learning and motivation: Advances in research and theory*, vol. 8 (pp. 47-89) . New York: Academic Press.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.

Rieman, J. Young, R.M., & Howes, A. (1996). A field study of exploratory learning strategies. *ACM Transactions on Computer-Human Interaction*, 3, 189-218.

Bachman, Lyle & Cohen, Andrew (1998) . *Interfaces Between Second Language Acquisition and Language Testing Research***. UK: Cambridge University Press, xi + 202.**

By Celso Henrique Soufen Tumolo

The book *Interfaces Between Second Language Acquisition and Language Testing Research* was edited with the concern of bringing together various researchers who have tried to overcome the separation of the two areas, SLA and LT, by raising and discussing relevant issues related to both. The result is a very important piece of writing. It is very dense reading, which presupposes, on the part of the reader, previous profound knowledge in both areas.

The various authors were requested to be based on the same article with title *Language testing – SLA research interfaces*, written by Lyle Bachman in

1988, for "a common point of departure", as the editors comment in the preface, which is included as an appendix in the book. It will be, thus, the first to be reviewed, breaking the order of the book.

In the article, Lyle Bachman criticizes the separation of SLA and LT as distinct areas of inquiry, SLA focusing on description and explanation of proficiency development, concerned about factors and processes that are part of or affects acquisition, i.e., the antecedents of proficiency, and collecting data from the analysis of learners interlanguage utterances, whereas LT attempting to describe language proficiency at a given stage, focusing on the results of acquisition, and collecting data from *ex post facto* correlational methods.

The author points to some changes, in the 1980s, since LT researchers were beginning to consider the nature of language proficiency to improve its processes, and to develop more useful measures, including the new views about communicative competence and communicative language use, resulting in a multicomponential view of language.

As part of these changes, LT researchers were becoming aware of the need to examine more closely the processes involved in testing, while SLA researchers were exploring approaches to assess language ability based on SLA developmental considerations. In the late 1980s, research into the nature of language proficiency and test development was already informed by theories of communicative competence and communicative language use.

To the author, it was the beginning of the interface between the two formerly distinct areas of LT and SLA, seen as complementary approaches to the inquiry into the nature of language proficiency and its acquisition. Three interfaces are presented in the article: 1) the covariance structure analysis of *ex post facto* correlational design; 2) the qualitative investigation of test taking processes; and 3) the development of L2 assessment instruments based on developmental sequences in L2 acquisition.

The author brings up a distinction between development and variation made by Pienenann and Johnston (1987), where development is seen as regular and predictable, and variation as the result of individual differences. Following this distinction, the result of a proficiency test cannot be an indicator of development in that accuracy will vary as a "function of both regular developmental sequence and individual variation across that sequence" (p. 190).

To tackle this problem, the author suggests a practical method of collecting information about the learner's SLA stage of development, i.e., the use of profile analysis, procedure which is "promising, since it addresses the need to consider acquisition history in the assessment of L2 proficiency" (p. 190).

The importance of articles such as this is that they challenge embedded views of language researchers and testers, and

more importantly, the view of practitioners in a foreign language class. Language researchers can understand the importance of the interface to keep on carrying out their studies. Language testers can learn how to construct tests taking into consideration the distinction presented between development and variation. Practitioners may have an informed decision to adopt, for example, the profile analysis, since they have been the ones who, ultimately, make most decisions concerning language testing of their students.

Chapter 1, Language testing – SLA interfaces: An update, written by the editors Lyle Bachman and Andrew Cohen, is an attempt for an update of the article included as the appendix. It aims at presenting a conceptual framework and approach for strengthening the interfaces between LT and SLA research. The authors claim that the foci of SLA and LT have not changed much since 1988, and present a chart in which they characterize the different foci, in terms of research perspective, focus, goals, and methodology.

Despite the differences in foci, they see a common area of interest as expressed in a central issue of describing and explaining variability in language acquisition (SLA) and test performance (LT). Accepting an important distinction put forth by Tarone (same volume), in which individual difference refers to differences in performance across individuals, and variation to shifts within the performance of the same individual, the authors focus on, and discuss the three sources of variability central to both fields: 1) individual differences in the language abilities that are acquired or measured; 2) individual differences in the strategies and other processes that individuals employ in language use, as well as in language test tasks and SLA elicitation tasks; and 3) variation in the tasks and context and their effects on language use, as well as on performance in language test tasks and SLA elicitation tasks.

These sources of variability are discussed by the authors in three sessions with many sub-sessions. The sessions are: 1) the nature of the language abilities acquired or measured; 2) strategies and processes in LT and SLA; and 3) variations in tasks and context and their effects on language use.

Although all points raised and discussed are extremely important for theoreticians as well as for practitioners, it is in the session *Directions for Future Research* that they bring to bear, in my view, the most important issue: the relationship of authenticity and the nature of language tasks.

Considering the relationship, and the claim, by the language testers, that test tasks have to be designed in a way that correspond to non-test language use tasks so as to allow for inferences about language ability, the authors present two dilemmas for LT and SLA: the observer's paradox and the bandwidth fidelity dilemma.

The first dilemma refers to the fact that the very act of observation "may change, in fundamental ways, the nature of language use that is being observed" (p. 22), implying that it is almost impossible to observe authentic language use and collect instances of language through LT test tasks or SLA elicitation tasks. The second dilemma refers to the reliability of the authentic language tasks, since broad authentic samples of language use might result in more generalizable but less accurate inferences of components of language ability.

In order to solve these two dilemmas, the authors propose a reconceptualization of authenticity as a relative quality, and cite the study by Bachman and Palmer (1996) as a suggestion for an approach to characterize and design test tasks that correspond to non-test language use tasks, that is, to "test takers' language use in nontest situations" (p. 23).

In chapter 2, Construct definition and validity inquiry in SLA research, Carol Chapelle presents a discussion on the SLA elicitation devices based on two principles - construct definition and validation, justifying her concern since the performance observed in these devices is used to "make inferences extending beyond the observed performance" (p. 32).

In examining the nature of construct definition, the author defines construct as a meaningful interpretation of the observed behavior, resulting from the inferences made based on the test results. The whole process going through inferences, interpretation, and definition of a construct requires performance consistency, which she discusses through three different perspectives: trait, behaviorist and interactionist.

The author points out that researchers are increasingly recognizing that understanding how traits and contexts interact will be more helpful in learning about language development and use. An inter-actionist perspective is, therefore, receiving strong support for a construct definition, since it specifies how language traits are put into use in context. Performance consistency is, thus, attributed to learner characteristics and the values of the contextual variables.

As for validity inquiry, the author defines validation as an "ongoing process of justifying particular interpretations and uses of test results" (p. 33). Justification includes relevance, utility, consequences, and specifically for testing, it implies the clarification of the value implications attached to the nature of the construct definition each test reflects and associated with particular test interpretation and uses whose objective is "to assess consciously the values implied by the choice of a test for a particular purpose" (p. 50).

The author stresses that tests need to be subjected to the process of validity inquiry to reveal what aspects are underlying the relevant signs and samples of learners performance, and that performance consistency should be the observable and observed data, and

adds that understanding the nature of operational settings across which consistent performance can be observed is necessary for a better idea of the interactionalist construct definition.

In chapter 3, Research on interlanguage variation: Implications for language testing, Elaine Tarone discusses the issue of interlanguage variation, grammatical accuracy and fluency, in response to the contextual variations. The author points out that the focus of SLA researchers on the study of authentic, meaningful language use differs from that of most SL testers, since the latter group relies on decontextualized language, and that the use of authentic and contextualized tasks for language testing has confronted with the problem of contextual variables causing systematic interlanguage variation. In this respect, she sets out to discuss interlanguage variation, the contextual variables influencing interlanguage variation, and the implication for LT.

Her claim, based on her review of research, that learner language is affected by language use context, and that a single number descriptive of language proficiency does not consider context changes, poses language testers the complex task of finding appropriate solutions for the problems presented.

In chapter 4, Strategies and processes in test taking and SLA, Andrew Cohen raises the issue of, and discusses, the relationship between the characteristics of test tasks and the strategies used by test takers. The author claims that very little attention has been paid to the processes used by the respondents when taking a language test to indicate their proficiency in reading, which, to him, may lead to misleading results because of the many "test-wise techniques that readers have developed for obtaining correct answers on such tests without fully or even partially understanding the text" (p. 91), which may be considered compensatory strategies, ranging from omission of material or production of different material, lexical avoidance, simplification, approximation, etc.

The author questions the reliability and validity of the commonly used testing instruments, since, in situations of testing, students are influenced by test-wiseness, that is, students are likely to use strategies not used under non-testing situations. In this sense, the author stresses the importance of SLA research in terms of providing fundamental information to the LT constructors for the identification of the strategies used by respondents which will "prove beneficial at all points in constructing, administering and interpreting language tests" (p. 108).

In chapter 5, Describing language development? Rating scales and SLA, Geoff Brindley contributes with a very important discussion on the validity of behavioral rating scales - a series of descriptions of stages or ranges of language behavior used to measure the learners underlying competence - usually used by language testers.

Many problems were discussed concerning behavioral rating scale, such as a) they are based on performance which varies since proficiency is a function of the processing skills required by the task, implying that different tasks call upon different language skills; b) they can only assess the present state of the language proficiency based on the interpretation of the learner's performance compared against a scale of expected behaviors, therefore, not assessing the state of the learner's interlanguage development, that is, the learner's developmental stage.

Brindley's contribution with his article adds more support to the idea of giving up the use of single unitary scales in favor of more qualitative instruments, which, in his opinion, can capture, to a greater extent, the complexity of the language learning process.

In chapter 6, *testing methods in context-based second language research*, Dan Douglas presents the existing debate among linguists as to the definition of context, as well as language proficiency, which makes clear that understanding the nature of context and what constitutes it, and the nature of language knowledge, is still a major problem.

According to the author, although it is accepted that context plays a role in language choice and acquisition, and that language proficiency is multicomponential, it is still unclear "what those components may be and how they interact in actual language use" (p. 142) in a given context, interaction which results in the construct of an internal context based on the internal interpretation of the inter-actants, viewed, therefore, as a "cognitive construct created by language users for the interpretation and production of language" (p. 146). A problem for testing methods that may arise is that the test takers may be interpreting the items in a different way from that intended by the test constructor, i.e., with a possible mismatch of interpretation between language tester and language taker.

Thus, language testers need to operationalize their notion of context in their tests, since the test environment – personnel, physical conditions, time, organization, instructions, level of precision, propositional content, and so on – leads the learner to perceive and assess the communicative situation based on which their interlanguage is produced on a test. The idea, concludes the author, is to capitalize on the test effects, providing "information interpretable as evidence of communicative competence in context" (p. 153).

In chapter 7, How can language testing and SLA benefit from each other? The case of discourse, Elana Shohamy presents and discusses, using the case of discourse, first, the contributions of LT to SLA, and secondly, the contributions of SLA to LT.

Concerning the contribution of LT to SLA, the author focuses on three main areas: a) defining the construct of language ability; b) applying findings

from LT to confirm and/or test SLA hypotheses; and c) providing SLA researchers with quality criteria for tests and tasks. The contribution of SLA to LT are in: a) identifying the language components that need to be elicited; b) proposing innovative tasks that can be used for language assessment; and c) informing language testers about language variations based on which to construct tests.

The author emphasizes that discourse analysis would benefit from the cooperation between SLA researchers and language testers, through which the former group would devise hypotheses and questions to be considered by language testers, resulting in new or revised hypotheses.

Thus, in this chapter, the author contributes with a clear explanation as the possible interfaces of the two previously separate areas, considering specifically language in use, i.e., discourse, and proposes a very interesting rating scale for discourse, based on which we can, through more qualitative analysis, have a more profound description of a learner language.

The book brings many important aspects for mutual contribution between LT and SLA. The interfaces between the two areas show the complexity of the processes involved. Books, such as this, have been published as attempts to make these processes clear, as well as to help in the construction of more reliable and valid testing instruments for more reliable and valid constructs based on the data collected. The dice is cast.