

## 8

### Discussion

The objective of this study was determining if the modifications made upon ASQ-BR (2010) in order to provide adjustments for the following application in 2011 improved its psychometrics characteristics. Aside from deepening psychometric theory, this dissertation aimed at discussing how modifying selected items could empirically change data collection. For that, a few psychological theories will be discussed as well as ASQ-BR's social and cultural adequacy to its real context of application (Borsa, Damásio, & Bandeira, 2012) and how both aspects could be the basis for the actual results.

Despite our discussion was organized in the following order: (1) descriptive statistics of the score, (2) inferential statistics of both application years, (3) dimensionality, (4) reliability analysis and (5) item analysis, we believe that the discussion would be clearer if scales were divided by domain. Scales should be understood as a whole and not for each result since inferential results can be influenced by data reliability and dimensionality. Besides, a modified item may start to contribute more – or less – to the total of the scale causing a change to the whole structure of the scale, *e.g.*, increasing or decreasing the scores of the children. We believe in this case it is better to discuss the results by domain, respecting age intervals. Despite doing that, this study does not try to answer why Personal/Social domain still presents two dimensions nor worries about items that have not been modified though showed significant differences along the two application years – what might have influenced the measure consistency (Cohen & Swerdlik, 2009). The limitations of the study and future directions for ASQ-BR will be presented at the end of this dissertation and will entail new suggestions for further research.

## 8.1 Communication

A baby's ability to communicate appears with its first unintended sounds and movements (Wagner, 2006). Babies' behaviors are unintentional in the beginning of life since infants do not expect any particular outcome from their behavior. Intentional communication is only established much later, when behavior becomes responses to particular stimuli or situations. Piaget (1953) suggests that in the first year of life babies tend to have basic emotional behaviors – crying, laughing and active looking – and its communications skills are associated to an attempt to imitate the world around him, copying adults' behaviors, joining and sharing attention with adults and peers and anticipating his/her own behaviors (engaging in games such as *where's mommy?*). After 9 months, other behaviors – babbling, vocalizing, hearing and one-word speech – lead to the development of a limited type of speech language. It is expected that children complexify their semantic and pragmatic understanding after 12 months of age until reaching adult-like speech at the age of five (Bialystok, 1986; Kidd & Bavin, 2002). The definition of communication behaviors in the ASQ-3-BR goes as follow:

“Ability of babble, vocalize, speak, hear and understand. Structure and express some thought so that his/her interlocutor understands it. Simple verbal structures, some speech complexity and correct use of plurals, complex and conditional verbal tenses.”

(Squires et al., 2009)

Several assumptions are possible regarding ASQ-3-BR items if the above mentioned assumptions by Squires et al. (2009) are considered. The first is that ASQ's questions consider more than language as communication, including any features of social interaction in any way other person understands. Engaging in joint attention behaviors and concerning about others perspective are important parts of communication skills (Wagner, 2006). Nevertheless, not only domain

Communication deals with the latter skills in ASQ. Domain Personal/Social also measures that. This dissertation is going to discuss the implications of having social skills measured in two different scales. Perhaps this is one of the possible explanations for consistent psychometric problems in the Personal/Social scale—both in 2010 and 2011 results.

In Communication, only two changes were empirically examined. In item *“O bebê fala três palavras como, por exemplo, “Mamã”, “Papá” e “Dá” (Uma “palavra” é um som que o bebê fala regularmente referindo-se a alguém ou a alguma coisa)”* changes were basically two: the inclusion of the word “Dá” (give) and the explanation of the item in parenthesis. The item clearly addresses the issue of one-word speech in the correct period of childhood development. These changes did not help empirically improve the item itself nor the scale, however it made the item more adequate to the evaluation setting, *i.e.*, child daycare centers. We thus considered those modifications to be more suitable to the environment where the assessment takes place, according to Borsa (2012).

The other modified item was *“Se você aponta para figuras e pergunta à criança “O que é isso?”, ela nomeia corretamente pelo menos uma figura? (Exemplos de figuras: bola, gato, carro, casa, etc.)”*, modified by including only the explanatory parenthesis. It addresses the child’s ability to name an object, in other words, the child’s semantic knowledge. The item was altered in three different age intervals – 20, 22 and 30 months – but only the last age interval showed statistical difference between years of assessment ( $\Delta=-0.02$ ). However, since an item contributes to the scale’s total score (Cohen & Swerdlik, 2009), when an item is changed, it is expected that its relation to the other items and to the scale as a whole is also empirically modified. Indeed, all Cronbach’s alphas of scales in age intervals 20, 22 and 30 months showed statistical difference, showing that the item modification, even when the item-total correlation was not significant, probably contributed somehow to alteration in the scale’s internal consistency. Empirically speaking, though, item modification was unsuccessful in two scales – 22 and 30 months – but successful in the 20-month communication scale. Despite that, item modification seems to not have influenced children’s performance in those scales, once no statistical difference was found between averages comparing years of assessment.

Semantic knowledge is expected to increase in different rates, developing in logistic progression (Kidd & Bavin, 2002), that is, there is rapid acquisition of meaning in early stages and progressive changing from quantity to quality as years go by. Probably, teachers' expectations for 20-month old children are different from expectations for older toddlers, which could lead to the difference in their evaluation.

In order to decide if those modifications should be incorporated to the ASQ-BR, four results were considered: (1) adequacy of the adaptation to child daycare centers, (2) Cronbach's alpha, (3) item-to-total correlation and (4) stability or improvement of children's performance in a given scale. In Communication scales, items were more adequate to its assessment context, no Cronbach's alpha was found below the established criteria ( $\alpha < 0.65$ ) and stability of children's performance was achieved in all communication scales. So, we recommend that the modifications implemented for ASQ-BR-2011 remain for further research with the ASQ-BR.

## 8.2

### Gross Motor Domain

Among ASQ-BR scales the Gross Motor domain was the most reliable according to Filgueiras (2011). Gross Motor abilities are those defined by controlling and coordinating movements in space with upper and lower parts of the body (J. Squires, Bricker, Twonbly, & Potter, 2009). They are indeed the first ones to develop in the human baby. Rapidly, baby suction reflex is followed by moving arms and holding the mother's breast to provide more security during breast feeding; movements of arms and legs get quickly under control to stiffen the body when the baby is picked up, culminating in walking before one year of age (Piaget, 1953).

“Broad bodily movements. Move arms to complete simple tasks as throwing an object or leaning against walls or handrails. Leg and feet coordination for balance and moving.”

(Squires et al., 2009)

Gross motor control, in terms of assessment, is a very straightforward construct in ASQ. Locomotion and coordination between legs and arms are basically the essence of the questions in this domain. Probably the simplicity of the construct is the reason for the better assessment: “(...) assessing a simple psychological construct is perhaps the best way to assure its validity and consistency.” (Clark & Watson, 1995). Because of its simplicity, modifications of items focused solely in adding a small assertion in parenthesis to the items. The three modified items had the same explanatory parenthesis: “(*Você pode observar isso na creche, no parquinho, em casa ou numa loja.*)” – the added term in bold. Few or no statistical differences between years of assessment were expected. However, one item showed significant decrease of item-total correlation leading to decrease in internal consistency measured by Cronbach’s alpha, while another item had the opposite behavior, improving both item-total correlation and the alpha of the scale. Other than that, several scales, regardless of the modifications, also showed Cronbach’s alpha increase or decrease.

Using the criteria above-mentioned to decide if the changes should remain, item-total correlation and Cronbach’s alpha did not changed enough to consider further modifications in those items. Adequacy was also better because the scale considers gross motor related behaviors at the very child daycare centers. Finally, significant differences between score averages between 2010 and 2011 occurred only in the 10-month age interval, thus the modifications seem to have not influenced children’s performance in the scale: a point in favor of the scale’s stability. Therefore, we recommend that changes in items in Gross Motor Domain remain. Among all the scales, only the 22-month age interval presented alpha below 0.65 and we suggest further studies to improve the scale’s reliability.

### 8.3

#### Fine Motor Domain

Fine motor skills are fundamental to children development. Countless evidences in different areas of cognitive sciences show that fine motor skills such as pointing, grasping and clinging can actually predict the emergence and development of other cognitive domains (James, 2010; Nieder & Dehaene, 2009; Pulvermüller, Shtyrov, & Hauk, 2009; Wilson, 2002). For example, Nieder and Dehaene (2009) suggest that pointing using the index finger is one of the first movements towards voluntary joint attention between baby and mother. This behavior is ultimately leading to communication. In ASQ fine motor scales, several aspects are measured from fine control of the finger movements to enough control to use scissors or a pencil adequately. The definition of fine motor control can be seen below:

“Movement and coordination of fingers and fingertips ability to use tools like knobs, scissors, taps, pencils and pens”.

(Squires et al., 2009)

Basically, items require children to execute a task, which can be problematic in child daycare centers because it may mean getting out of the normal path of daily activities. One of the main concerns of teachers during the pilot study was exactly the ability and the adequate training of professionals involved in the assessment to understand the correct way to ask children to do any of those tasks and respond to them in a homogeneous way.

Among Fine Motor items, three were altered. Two of those items clearly improved the scale. Item “*A criança liga e desliga interruptores de luz? Abre a maçaneta ou trinco da porta? Abre e fecha torneiras? Caso a criança faça uma dessas atividades marque sim.*” initially only required the children to show if they knew how to switch the lights on or off. By adding other fine motor behaviors, the scale improved homogeneity in the 27-months scale. This probably happened because the modification allowed teachers and caregivers to observe other behaviors leading to more consistent responses.

The other improved item in Fine Motor scales was “Faça uma linha dividindo ao meio uma folha de papel. Usando tesoura sem ponta, a criança corta o papel ao meio, mais ou menos em linha reta, fazendo com que as lâminas se

abram e se fechem? (*Observe cuidadosamente o uso da tesoura por razões de segurança.*)”. The item was changed to include a parenthesis guiding teachers and caregivers to be careful watching the children when using scissors. The message served for two purposes: (1) for safety reasons, and (2) to make clear that the child had to be tested to answer this item. During the pilot study, teachers reported that several caregivers were answering the ASQ-BR without testing children in all the required behaviors, but only in a few. This could be misleading and jeopardize statistical results. Analyzing item statistics and internal consistency, we strongly recommend that those changes remain in the ASQ-BR.

Regardless of that, an unexpected phenomenon was observed: seven age intervals showed improvement in children’s performance in fine motor skills. There are several possible hypotheses to explain and they will be addressed ahead in this study.

## 8.4

### Problem Solving Domain

Problem solving is a controversial construct because it is focused on the outcome – the behavior itself, *i.e.*, if the child solves a problem – and not on the process. Cognitive science is more concerned, nowadays, with the processes leading to an answer than with the answer itself (Munakata, Casey, & Diamond, 2004). Nevertheless, research on information processing in babies and toddlers is still very incipient (Hackman & Farah, 2009). Thus, instrumentalizing the construct would be more coherent with the literature than assessing just the outcome. Problem solving can be defined in several different ways, thus we are going to stick to Squires’ et al. (2009) definition:

“Respond appropriately to external and internal demands of the environment, such as: taking an object from inside another, handling two pieces of information at the same time, imitating or copying adults, attributing meaning, recognizing and categorizing objects and people.”

(Squires et al., 2009)

Among the seven modified items, four showed actual statistical differences between years of assessment. Two of those items were in the 10-month scale.

Both item-total correlation and Cronbach's alpha improved significantly, what confirmed the implementation of changes.

One item had low item-total correlation and its change led to a decrease of internal consistency. Item “*Enquanto a criança observa, alinhe quatro objetos, como blocos ou carrinhos, em uma fileira, como se fosse um trenzinho. A criança copia ou imita você e também alinha quatro objetos em uma fileira? (Você também pode usar carretéis de linha, caixinhas ou outros brinquedos.)*” had parentheses included so teachers and caregivers had more tools to observe the behavior. However, observing more children behaviors in this case was misleading at least for children in the 27-month interval. Average scores of the scale were statistically higher in 2011 than in 2010. That can suggest that the modification actually made the scale easier, probably with an important contribution of the item. A specific study of item difficulty would probably help understand what happened with the item. So, regarding this item, we recommend to change it back to its previous form until changes are reengineered for further studies.

The last modified item, “*Se a criança quer alguma coisa que não consegue alcançar, ela procura alguma coisa para subir e alcançar o objeto (por exemplo, para pegar um brinquedo sobre uma prateleira ela sobre no bloco de espuma)?*” had an interesting behavior regarding its item-total correlation. In the 30-month scale, no statistical difference was found. In the 33-month interval a significant decrease was observed while the complete opposite happened with the 36-month interval, which showed significant improvement. One can hypothesize that there is some age effect in the item. But, after evaluating children's performance on the three scales, no significant difference was found. We thus believe that the modifications only helped adapt the item in the assessment context. Based on that, we recommend that six of the seven items remain altered. Five scales, though, should be studied in future research because of impaired internal consistency ( $\alpha < 0.65$ ) in intervals 20, 22, 24, 27 and 54-month.



## 8.5

### Personal/Social Domain

This domain is indeed controversial. Filgueiras (2011) and Filgueiras et al. (2013) argue that the low homogeneity in this scale, measured by Cronbach's alpha, is due to the scale's attempt to measure two different constructs at the same time. Indeed, the description of the domain according to Squires is:

“Ability to be independent and relate to other children and adults. Verify if: the child looks for help when needs something, is able to engage in relationships with other people, can identify with elements socially established for his/her individuation, is independent in daily tasks like eating, getting dressed, and clean him/herself.”

(Squires et al., 2009)

The ability of being independent in a child – personal skills – is associated with behaviors of autonomy, *e.g.*, not paying attention to others, but at his/her own will (Kochanska, Murray, & Harlan, 2000). In contrast, social skills are the ability to interact appropriately, concerning with others and engaging in pleasant interactions for both parts (Jane Squires, Bricker, & Twombly, 2004). Thus, modifications in this scale should benefit from concern in measuring both constructs in a unidimensional fashion. The authors of the scale have actually attested that it lacked some precision statistically and, for that, created ASQ: Social Emotional (ASQ:SE) for dealing with the issue (J Squires, Bricker, Heo, & Twombly, 2010).

Regarding dimensionality, based on the latter argument, two dimensions would be expected. However, according to Filgueiras et al. (2013), personal and social skills could be “two faces of the same coin”. Their study conducted several factor analysis using different techniques and only three among the twenty analyzed scales presented bidimensionality. In the present study, only one of the scales shows such results, the 60-months scale. Thus, scales can be considered unidimensional.

With respect to internal consistency, the present study gave one step towards a better scale. In 2011, only seven scales had alphas below the established criteria, against twelve in 2010. However, it seems that even with the impressive

improvement of the scales – both in internal consistency and item-to-total correlation – problems with Personal/Social scales still remain. Future studies are therefore needed to carefully analyze Personal/Social scales.

## 8.6

### Limitations of the study

The results found in the present study are not entirely comparable to the ASQ-BR used in 2010. Several variables were altered between years of assessment, when conclusions derived from the present study were not available. The first variable changed between 2010 and 2011 was the inclusion of preschool children in the sample, which remarkably increased the sample size of the 60-month age interval.

The second variable that probably influenced results were questionnaire completion by teachers recently hired (in 2011) who had not been trained on ASQ-BR. Actually, the Secretary of Education of Rio de Janeiro hired 1.500 teachers for child daycare centers through a public selection (Rio de Janeiro, 2010) at the beginning of 2011. That meant an increase of over 10% of professionals using ASQ-BR-2011. Despite the efforts for adequate training given to the directors of municipal schools and daycare centers it is not possible to entirely guarantee the quality of the training passed on to the teachers by the directors. The novelty of the profession for the recently hired teachers as well as possible inconsistencies on directors' training may be reason for the average difference.

The third possible explanation for significant differences between scales is the development of new activities in child daycare centers in Rio de Janeiro in virtue of the 2010's assessment. The Secretary of Education had developed a free adaptation of the ASQ-3's book of learning activities (Rio de Janeiro, 2011). The book gives teachers and caregivers guidelines of how to improve children performance in classroom with respect to each ASQ domain. There are no official reports or information about the effects of those activities in municipal children enrolled in daycare centers in Rio de Janeiro but the initiative might be an explanation to the difference of averages between 2010 and 2011.