The IADL Profile: Development, content validity, intra- and interrater agreement

Carolina L. Bottari, Clément Dassa, Constant M. Rainville, Élisabeth Dutil

Abstract

**Background.** Occupational therapists are frequently asked to document the interplay between individuals’ neuropsychological deficits and the requirements of their daily lives. **Purpose.** The present study was designed to develop and validate a measure of independence in everyday functioning that considers recent advances in research regarding the ecological assessment of executive functions. **Method.** Experts (n=8) judged the IADL Profile’s content validity and the tool was pilot tested with the target population (n=8). To document the tool’s reliability, 30 patients aged 16 to 65 with moderate or severe TBI were recruited. A trained examiner administered the IADL Profile, and three trained judges rated video recordings on two occasions. **Results.** An eight-task (29-item) test was developed. Comparing ratings of four raters, 95 percent of kappa coefficients indicated moderate to almost perfect agreement, and 94% showed almost perfect intrarater agreement. **Implications.** The IADL Profile provides occupational therapists with a set of measures of IADL independence with strong preliminary evidence of reliability.

Résumé

**Contexte.** Des demandes sont souvent adressées aux ergothérapeutes pour documenter les liens existants entre les difficultés observées dans les tâches de la vie quotidienne et les troubles neuropsychologiques. **But.** Cette étude a été conçue dans le but de créer une mesure de l’indépendance dans les activités de la vie quotidienne et d’en démontrer la validité. Cette mesure tiendra compte des progrès récents réalisés dans le domaine de la recherche en ce qui a trait à l’évaluation écologique des fonctions exécutives. **Méthodologie.** Des experts (n=8) ont évalué la validité de contenu du Profil des Activités Instrumentales (PAI) et un pré-test a été complété auprès de la clientèle cible (n=8). Afin d’évaluer la fidélité de l’outil, trente personnes ayant subi un traumatisme crânien modéré ou sévère, âgées de 16 à 65 ans, ont été recrutées. Un examineur qualifié a administré le PAI et trois juges qualifiés ont analysé les bandes vidéo à deux reprises. **Résultats.** Un test composé de huit tâches incluant au total 29 items a été développé. En comparant les résultats des quatre évaluateurs, 95% des coefficients de kappa indiquent un accord modéré à presque parfait alors que 94% indiquent un accord intra-évaluateur presque parfait. **Conséquences.** Le PAI propose aux ergothérapeutes une série de mesures de l’indépendance dans les activités de la vie quotidienne dont la fidélité est démontrée.

Citation: Bottari, C. L., Dassa, C., Rainville, C. M., & Dutil, É. (2010). The IADL Profile: Development, content validity, intra- and interrater agreement. Canadian Journal of Occupational Therapy, 77, 90-100. doi: 10.2182/cjot.2010.77.2.5

Submitted: 10 June 2009; Final acceptance: 20 October 2009.

This research was supported by operating grants awarded to the authors by the Fonds de la recherche en santé du Québec, the Association québécoise d’établissements de santé et de services sociaux, the Association des établissements de réadaptation en déficience physique du Québec, the Société d’assurance automobile du Québec, the Québec Rehabilitation Research Network (grant no: 6623), and the Canadian Occupational Therapy Foundation. The principal investigator on the grant was Élisabeth Dutil.
Executive dysfunction is increasingly thought to have important consequences on an individual's independence in everyday activities (Godbout, Grenier, Braun, & Gagnon, 2005; Royall et al., 2007). Executive functions have been defined as the ability to plan and carry out complex goal-directed behavior (Lezak, 1989; Luria, 1966, 1973) and include such things as formulating goals, planning, carrying out tasks, and verifying that intended goals are attained.

A core element of occupational therapy is the assessment of independence in everyday activities. However, to our knowledge, limited attention has been given to re-examining performance-based measurement instruments of instrumental activities of daily living (IADL) to ensure that these measures capture the repercussions of related deficits in executive functions. Due to the central importance attributed to executive functions for independent living (Shalllice & Burgess, 1991), it is essential that these measures capture potential repercussions of executive function deficits and the complex interactions of potential disabilities with the many environments in which the person undertakes daily responsibilities.

Several performance-based IADL tests are thought to underestimate the repercussions of executive function deficits on IADL independence (Bottari & Dutil, 2000). Tests such as the Assessment of Motor and Process Skills (Fisher, 2003), the Arnadottir OT-ADL Neurobehavioral Evaluation (Arnadottir, 1990), and the Executive Function Performance Test (Baum et al., 2008) tend to overly structure clients and remove the possibility of observing the individual's ability to formulate goals and to plan, though both are components of executive functions that are of crucial importance to independent living. Moreover, few performance-based IADL measurement instruments consider the reality experienced by the person, taking into account the different environmental contexts in which he or she exercises his or her roles (Bottari, Dutil, Dassa, & Rainville, 2006). A recent review of performance-based personal activities of daily living (PADL) and IADL measurement instruments ranked the ADL Profile first among published standardized ADL measures with regards to its correspondence to constructs of importance to occupational therapists, such as holistic perspective and client centeredness (Klein, Barlow, & Hollis, 2008). Interestingly, this test was also designed with the intent of documenting the repercussions of executive function deficits on independence in everyday activities for individuals with severe traumatic brain injuries (Dutil, Bottari, Vanier, & Gaudreauult, 2005).

The ADL Profile, ideally administered within the person's home and community environment, uses a non-structured evaluation approach to maximize the possibility of observing deficits related to goal formulation and planning (i.e., executive function deficits). Though this has been summarized succinctly elsewhere (Bottari, Dassa, Rainville, & Dutil, 2009b), the semi-structured approach is particularly reflected in the instructions and graded cueing provided to the examinee. At the onset of the test, examinees are systematically informed that the examiner will do and say as little as possible throughout the duration of the test so as to fully observe their ability to manage on their own. They are provided general instructions that ask them to do that which they would normally do at this time of the day leaving ample opportunity for goal formulation and planning. Examinees are then given the opportunity to do ADL tasks of their choosing without cueing from the examiner (so as to allow a thorough observation of the client's ability regarding executive functions). In the event that examinees opt to do tasks that are not in the ADL Profile, they may simply be asked to consider an alternate goal of their choosing. When administering the test, the examiner does not intervene, even if the examinee makes errors, because observing the person's ability to monitor and correct these errors without assistance is of central importance to documenting independence in consideration of executive functions. Cueing is only provided if and when the examiner judges that the person, despite having had sufficient time to attempt the task on his or her own, nonetheless remains unable to do any one component of importance without help. Congruent with the ADL Profile's non-structured approach, the examiner can withhold cueing for extended periods of time (e.g., 5 to 10 minutes) unless, of course, the situation is judged as dangerous, whereupon more rapid intervention is provided. When assistance is deemed necessary, it is provided in a graded fashion, that is, from more open-ended general questioning to specific cueing.

The ADL Profile also provides a task analysis framework that considers executive functions as each ADL task is qualitatively analyzed and scored on the basis of the person's independence with regards to each of the following four operations: formulating a goal, planning, carrying out the task, and verifying attainment of the initial goal. Independence on each operation is scored using a four-point ordinal rating scale (3—indepen dent; 2—independent with difficulty; 1v—requires verbal assistance; 1p—requires physical assistance; 1vp—requires both verbal and physical assistance; 0—dependent). The task score is simply the lowest score given to any one operation related to the particular task. For instance, a person who scored 3 (independence) for formulating the goal of using public transportation, 1v (requires verbal assistance) for planning, 2 (independence with difficulty) for carrying out the task, and 1v (requires verbal assistance) for verifying attainment of the initial goal would be given a score of 1v (requires verbal assistance) for this task.

The ADL Profile has 20 PADL and IADL tasks, 17 of which are included in the performance-based test and 3 of which are exclusively documented via a semi-structured interview. One of the unique features of the ADL Profile is the complementary nature of the performance-based test and the questionnaire or semi-structured interview. Documenting independence from these complementary perspectives adds valuable information when interpreting test results and a better approximation of real-world functioning. For instance, documenting pre-traumatic ADL habits (frequency of practice, level of responsibility) via the ADL Profile interview (subject, significant other) informs the examiner of the contexts (familiar, novel) in which the difficulties are observed.

To offer further clarity to the goal of the ADL Profile, it may be important to contrast the features of this measure of ADL independence to those of another type of test, tests developed as ecological measures of executive functions. One such test is the Multiple Errands Test (MET) (Shalllice & Bur-
In this test the examiner documents the number and type of errors made by the subject while carrying out a series of predetermined tasks, that is, the subject is given a list of tasks to do and specific rules and constraints that need to be respected while carrying out the specified tasks. The MET provides valuable information regarding the subject's ability to plan within the context of a multitask situation. However, it does not allow for the observation of goal formulation nor of planning within an unstructured context in which the person does not have the benefit of a detailed list to follow, a common occurrence in everyday life. When administering the MET, the examiner does not intervene as the aim is not to verify level of independence or ability to attain a goal with assistance, but rather to ascertain the presence or absence of executive dysfunction based on the number and types of errors observed. In contrast, the ADL Profile seeks to document an individual's ability to manage everyday activities with or without help and to obtain valuable information for treatment interventions. Both task-related abilities and difficulties are documented, not only errors. Hence, in our view, the need for a measure of IADL independence that takes into consideration executive function deficits has not been met by recent work on the MET or other similar work on the ecological measure of executive dysfunction.

The present study was designed to expand upon previous work on the ADL Profile as we developed and validated an alternate version of the ADL Profile (performance-based test only), the IADL Profile, which considers recent advances in the field of executive functions, more specifically in the area of the ecological assessment of executive functions (Burgess et al., 2006). The IADL Profile is an ecological measure of independence administered in a subject's home and community environments that aims to establish whether the subject's main difficulties pertain to formulating goals, planning, carrying out the task, attaining the initial task goal, or a combination of these, all important components of executive functions that can be impaired subsequent to a traumatic brain injury (TBI). Our primary objective was to develop a tool with a greater degree of task complexity (e.g., complex sequence of six interrelated tasks linked to the overarching goal of receiving guests for a meal), fewer tasks to increase the feasibility of administering the tool within a single session, increased consideration of the complex interactions between environmental demands and the person's abilities, task definitions that consider components of executive functions, task analysis based on more explicitly defined operations related to executive functions, an improved rating scale, and improved scoring criteria. The target population was individuals aged 16 to 65 with moderate or severe TBI. Subjects were recruited from three post-acute care settings, including four to five individuals with mild, moderate, or severe TBI in their home and community environment; and we corroborated the initial task selection by a second review of the literature.

Phase 2: Tool construction
Phase 2, construction of the tool, consisted of four main steps: (1) developed a complex sequence of inter-related tasks related to the overarching goal of hosting a meal to increase overall task complexity and provide a greater opportunity to document the complex interactions between environmental demands and the person's abilities; (2) updated the task of "telephoning for information" from the ADL Profile to include more recent options, including use of the Internet; (3) revised the task "making a budget" from the ADL Profile to increase its overall level of complexity (yearly rather than monthly budget); and (4) elaborated the user's guide of the first prototype of the IADL Profile (version 1.0), including goal, evaluation approach, administration procedure (e.g., environment, equipment), scoring guidelines and procedure, and description of task-evaluation scenarios (definitions, administration context, instructions to person, and scoring guidelines). Tasks included in this first prototype were the following: doing daily house cleaning, putting on clothes and shoes, walking or moving outdoors, preparing a hot meal, having a meal, telephoning for information, managing one's finances and using public transportation.

Phase 3: Pilot testing
Phase three of the development of the IADL Profile (version 1.0) involved pilot testing the prototype with eight moderate and severe TBI. Subjects were recruited from three post-acute care settings.
TBI programs in the province of Quebec. Inclusion criteria were patient age between 16 and 65 years, moderate or severe TBI, and proficiency in French. The severity of the TBI was estimated with the Glasgow Coma Scale (GCS) (Teasdale & Jennett, 1974) and duration of posttraumatic amnesia. Exclusion criteria were disorientation, severe mobility deficits, and history of hospitalization for psychiatric disorders. Prior to the evaluation, subjects were given verbal and written information on the study and signed statements of informed consent. Subjects were informed that they were free to withdraw from the study at any time. The study was approved by the ethical review boards of participating centers.

The tool was tested for feasibility (clarity of instructions, ease/ difficulty of administering the IADL Profile in a person’s home and community environment, applicability of selected tasks, acceptability of test for subjects in terms of non-structured approach, and administration time, ease of scoring, etc.) using the examiner’s observations, logged in a research journal and subsequently analysed. All subjects were evaluated with the IADL Profile in their home and community environments by the primary investigator. She was not involved clinically with the clients, whom she met solely for testing purposes.

Phase 4: Validation

Content validity study.

To examine the tool’s content validity, the user’s guide to the IADL Profile (version 1.0) was submitted to an international committee of experts identified by university professors as having the required expertise either in executive functions and/or IADL, or in instrument development and validation. Experts were contacted via e-mail and invited to participate in the study. They were also invited to propose the names of other individuals judged to meet the study’s selection criteria. Selected content specialists (occupational therapists, neuropsychologists, neurololgists, and physiatrists) were required to have five years or more of research or clinical experience with individuals with a TBI or frontal tumours and recognized expertise in executive functions and IADL. At least one expert was required to be a specialist in research methodology and instrument development. Content specialists were asked to judge the pertinence and clarity of the following aspects of the tool: task definitions, instructions given to the person by the examiner, definitions of operations underlying task analysis, and rating scale. Specialists in research methodology and instrument development were only asked to judge the clarity of the aforementioned aspects of the tool. A three-point rating scale was used: (1) not pertinent/not clear; (2) more or less pertinent/more or less clear; (3) pertinent/clear. Both quantitative (frequencies) and qualitative (experts’ comments) analyses were completed.

Data were reviewed by our research team with particular attention given to items that one or more experts judged either not clear or not pertinent. Items judged to be more or less pertinent or relevant were also carefully reviewed. Based on these findings, required tool modifications were decided upon by consensus of our research team using the theoretical framework and goals of the tool as a guide.

Reliability study.

Reliability is defined as the consistency of measurement over time or the precision of the rater’s measurement (Crocker & Algina, 1986). Reliability refers to the extent to which the measure yields the same score each time it is administered, all other things being equal (i.e., no true change in the attribute being measured has occurred). In this phase of the study, the final prototype of the IADL Profile (version 2.0) was submitted to a reliability study of the intra- and interrater agreement. The IADL Profile (version 2.0) consists of 29 items scored by a rater, that is, 29 scores for five tasks with four operations and three tasks with three operations each, and is rated using a five-level ordinal scale of independence. Non-observed behaviors or actions are rated 8 (not observed for reasons intrinsic to the person) or 9 (not observed for reasons extrinsic to the person). The final selection of six tasks linked to the overarching goal of preparing to host a meal for unexpected guests was labeled as follows: dressing to go outdoors, going to the grocery store, shopping for food, preparing a hot meal for guests, having a meal with guests, and cleaning up after the meal. The final two tasks (obtaining information, making a budget) are complex single tasks. The latter are rated on only three operations: goal formulation operation is not rated as it is the examiner who formulates these goals. Also, the task of having a meal is only scored on three operations (planning, carrying out task, and verifying attainment of initial goal). The operation related to formulating a goal had to be removed from the test as it was shown to lack variance (all subjects were rated as independent) in a previous factorial validity study (Bottari et al., 2009b). The test is administered within the person’s home and community environments.

Participants.

Subjects were recruited from eight post-acute TBI programs in Quebec. Inclusion criteria were patient age between 16 and 65 years, moderate or severe TBI, and proficiency in French. The severity of the TBI was estimated with the Glasgow Coma Scale (GCS) (Teasdale & Jennett, 1974) and duration of posttraumatic amnesia. Exclusion criteria were disorientation, severe mobility deficits, and history of hospitalization for psychiatric disorders. The sample size was limited by practical factors whereby it was determined that a maximum of 30 subjects meeting our selection criteria could feasibly be recruited in the time available. Based on the procedure developed by Donner and Eliasziw (1992), which is “accurate in samples involving as few as 25 subjects” (p. 1512), it was expected that a sample of 30 would yield a power of 80% to detect a kappa larger than 0.45 under the null hypothesis of kappa = 0 with a significance level of 5% . Prior to the evaluation, subjects were given verbal and written information on the study and signed statements of informed consent. Subjects were informed that they were free to withdraw from the study at any time. The study was approved by the ethical review boards of the participating centers.

Raters.

The four occupational therapists who participated as raters in this study had been practicing occupational therapy for a mean of 5.25 years (range 1–11) and had been working in TBI
programs for a mean of 4 years (range 1 – 8). Raters underwent a standardized training regimen, a three-day training session given by the primary investigator (R4), also first author of the IADL Profile. Here raters were introduced to the architecture of the test (objectives, conceptual frameworks, variables, administration procedure, scoring, and interpretation) and practiced task analysis and scoring supported by videotaped administrations of the tool with TBI patients. Raters were also provided with the procedure manual for the IADL Profile, which contains specific scoring guidelines.

All subjects were evaluated with the IADL Profile by the primary investigator (R4) in their home and community environments. The evaluation was recorded with a video camera by a research assistant and subsequently analyzed by the three other raters (R1 – R3) who completed independent ratings of the evaluations. Three raters viewed videos (R1 – R3) of the test on two occasions (t1, t2). The two ratings of each video were separated by a one-month interval, time enough to reduce the effect of memory on the ratings.

Procedures.

Data collection took place between February 2005 and October 2005. The IADL test was administered on a single occasion and lasted approximately three hours. Repeated testing was not feasible as it would have been unreasonable to ask subjects with known problems of fatigue to repeat the evaluation, and having four raters in the subject’s home at any one time would have been overly intrusive and generally not possible. Hence, video recordings were used to allow raters to score the test on two separate occasions. Use of video recordings also minimized error associated with changes in IADL performance secondary to learning (particularly important in relation to executive functions, as tests are most sensitive to executive function deficits during initial testing when the test situation is most novel).

Data analysis.

To examine the reliability of the instrument, linearly weighted kappa statistics (κ) and percentage agreement (PA) were computed. Weighted kappa serves to establish the concordance between raters, correcting for concordance by chance on ordinal data (Landis & Koch, 1977) and weighting for the amplitude of discrepancies between the ratings. PA describes the number of occurrences in which raters are in agreement by chance or not.

The complete data set was analyzed with linearly weighted κ to identify items for which agreement was poor to fair between the three raters (R1 – R3) and the main evaluator (R4), whose rating was considered the norm. Percentage agreement (PA) was calculated on all items to complete the information. Analyses with kappas and PA were based on all known data (Landis & Koch, 1977) and weighting for the amplitude of discrepancies between the ratings. PA describes the number of occurrences in which raters are in agreement by chance or not.

The complete data set was analyzed with linearly weighted κ to identify items for which agreement was poor to fair between the three raters (R1 – R3) and the main evaluator (R4), whose rating was considered the norm. Percentage agreement (PA) was calculated on all items to complete the information. Analyses with kappas and PA were based on all known levels of the scale. The sample size was uniform for all items (n=28). Two subjects had to be removed from the sample due to missing data (unobserved operations). Kappa coefficient values between 0.81 and 1.00 represent almost perfect agreement beyond chance, values between 0.61 and 0.80 substantial agreement, values between 0.41 and 0.60 moderate agreement, values between 0.21 and 0.40 fair agreement, values between 0.00 and 0.20 slight agreement, and values below 0.00 poor agreement (Landis & Koch, 1977).

Results

Phase 1: Planning

Summary of literature reviewed.

Activities of daily living and traumatic brain injury.

Generally, at one, two, or several years after a TBI, studies have found that a high percentage of subjects continue to require assistance for ADL (more specifically IADL). Important deficits have particularly been noted with tasks such as meal preparation, shopping, housework, and personal finances (Dawson & Chipman, 1995). Indeed, as many as 90% of individuals with a severe TBI at 24 months post-trauma were shown to require assistance when using public transportation, running errands, preparing a hot meal, doing laundry, and doing daily house cleaning (Dutil et al., 1995). These findings were useful in identifying the IADL tasks that would be retained in the IADL Profile.

Executive functions and everyday activities.

To date, the type of everyday tasks thought to solicit executive functions remains controversial. Errors of action observed in individuals with an action disorganisation syndrome have been shown to occur during the performance of familiar multistep tasks, such as making a cup of tea (Schwartz, Mayer, FitzpatrickDeSalme, & Montgomery, 1993). Alternately, it has been suggested that executive functions are specifically required for novel and complex tasks (Shalice & Burgess, 1991). There is, however, a growing consensus regarding the use of ecological everyday tasks such as shopping and cooking to study executive functions, though not all tasks may be as sensitive to executive function deficits (Burgess et al., 2006; Chevignard et al., 2000). Based on these findings, we suggest that the research presented in the current study was warranted to increase the overall complexity level of the tasks of the ADL Profile.

A second element of importance is the testing approach. A critical problem underlying most tests is that the approach remains overly structured and thus fails to allow for the observation of errors related to certain key components of executive functions (e.g., goal formulation and planning) (Lezak, Howieson, & Loring, 2004). The ADL Profile is possibly the only performance-based ADL tool found in the literature to have operationalized a non-structured approach that permits the observation of goal formulation because subjects are simply asked to do as they would normally do and the evaluator observes the person’s ability to formulate goals, plan, and carry out activities. Studies carried out on over 100 severe TBI subjects have shown the viability of this approach and its acceptability for subjects (Dutil et al., 1996).

Though a number of error-coding systems have recently been developed (Rainville, Passini, & Marchand, 2001; Schwartz, Segal, Veramonti, Feraro, & Buxbaum, 2002; Shalice & Burgess, 1991), the ADL Profile is one of the very few measurement instruments to have developed a task analysis guide that permits the association between observable behaviours
and verbalisations and four components of executive functions (goal formulation, planning, execution, and attaining initial task goal). Formulating a goal is defined as the ability to find a solution to satisfy a need or solve a problem situation. Planning is defined as the capacity to think about the initial conditions before acting; identify alternatives; choose the most adequate alternative, and develop a general strategic and tactical plan of actions (sequence of actions or steps). Carrying out the task is defined as the capacity to initiate one's action plan; carry out the plan of action while adapting to errors or novel situations; perceive errors in planning and execution and adjust actions in relation to perceived errors and new or unforeseen situations. Verifying attainment of the initial goal is defined as the capacity to verify that the task that was initially planned was carried out; compare the final result to the initial goal; accept or reject the results; end the task or start the process again when the result is not attained. The task analysis framework of the ADL Profile examines the overall process of task realisation with a particular interest in documenting both the strategies used by the person to accomplish the task and to identify the stage at which breakdown in executive behaviour occurs. Application of the task analysis guide has, however, proven to be challenging for clinicians and has shown a need for further development and refinement. Based on these findings, we proposed to develop a new measurement instrument that would build on the strengths of the ADL Profile and attempt to temper some of its weaknesses.

Phase 2: Tool Construction

A core aspect of tool construction involved the development of a complex sequence of inter-related tasks related to the overarching goal of hosting a meal for unexpected guests. Initially, five tasks were included in this sequence: putting on clothes and shoes, walking or moving outdoors, shopping, preparing a hot meal, and having a meal. In order to achieve the intended goal of creating an evaluation scenario that closely approximated the requirements of a complex everyday life situation that explicitly tapped into executive functions, we framed the evaluation context so as to require the simultaneous planning of the full series of embedded tasks necessary to attain the ultimate goal of hosting a meal. This was done by giving examinees a single very minimally structured instruction for the complete series of tasks: examinees were informed that they had unknowingly invited guests for lunch, and that they should prepare to receive them. No equipment or ingredients were provided and no other form of assistance was given (e.g., examinees were not brought to the grocery store) unless deemed necessary.

Tool construction also involved a revamping of the scoring system of the ADL Profile. First, the assistance score of the ADL Profile was a single score that consisted of verbal assistance, physical assistance, or both verbal and physical assistance. This definition of the assistance score was considered to include an overly broad range of subject abilities. We therefore initially modified the four-level ordinal scale to a six-level ordinal scale (independence, independence with difficulty, need for physical assistance, need for verbal assistance, need for physical and verbal assistance, dependence), splitting the assistance score into three separate ratings. Second, in order to further operationalize executive functions within the IADL Profile, the description of the rating scale as applied to each of the four operations (goal formulation, planning, carrying out, and attaining goal) was expanded to include examples of behaviours and amount and type of assistance offered to illustrate functioning at all levels of independence on each operation.

Phase 3: Pilot Testing

The average age of the subjects was 42.9 years. There were an equal number of men and women, and the average time post-injury was 5.8 months. Five subjects had severe TBI and 3 had moderate TBI. Average duration of posttraumatic amnesia was 28.8 days. Average education was 12.9 years. Six were living in their own homes, one was hospitalized within an inpatient rehabilitation centre, and one was living in a transitional living apartment for people with a TBI. Six walked with or without an orthosis for outdoor mobility and two required the use of a cane.

Overall, subjects responded well to the test in terms of its non-structured approach and its administration within their home and community environments. Yet, the need for a number of modifications to the prototyped was still identified. For instance, the duration of the test was deemed too long for most TBI subjects to carry out during a single session due to fatigue. Hence, the task of using public transportation was removed, not only due to the time required to administer the task but also due to the observed absence of public transportation in the vicinity of the homes of several subjects. Also, certain problems were identified with subjects' comprehension of the following initial unstructured instructions provided by the examiner:

You invited friends over for supper. I would like to observe what you would normally do to get ready. During the evaluation, I will simply observe how you manage. You can ask me questions if you like, but please try to do things as though you were alone. Can you please repeat what I just said? Do you have any questions regarding what I am asking you to do?

After hearing these instructions, certain subjects had difficulty understanding that the examiner's request implied actually preparing a meal and running errands and that it was not a hypothetical scenario to be responded to orally. Some subjects gave a detailed verbal description of what they would do if the situation presented itself but failed to understand that they were being asked to actually plan and carry out tasks related to the scenario.

To improve the subjects' comprehension of the request as a real situation, a general introduction was added. For example, the examiner says:

We would like to know how you manage in your everyday activities, that is, activities that you generally do inside and outside of your home. More specifically, we would like to know, following your accident, if any changes have occurred in your ability to carry out your everyday tasks.

The subject is informed that the examiner will assume the role of an observer, saying and doing as little as possible, so
that the evaluation more clearly reflects how he or she functions when alone. The examiner verifies the subject’s acceptance of the evaluation context and responds to any questions before proceeding with the presentation of more specific instructions such as:

Without knowing it, you invited my assistant and me to have lunch with you. Please get ready to receive us. We will assume any incurred expenses for a maximum of $20. Can you tell me in your own words what I have just explained to you? Would you agree to do this? Can you now tell me in your own words what you are going to do?

The final version of the instructions was incorporated into the new prototype (version 2.0).

**Phase 4: Validation**

**Content validity study results.**

Of the 14 selected experts, 8 responded to our questionnaire. Among the eight experts who responded, one was a specialist in research methodology and instrument development, and seven were content specialists (occupational therapists, neuropsychologists, neurologists, and physiatrists) with both clinical and research experience. Their research experience was on average 14 years (range of 5–23 years) and their clinical experience with TBI was on average 18 years (range of 9–23 years).

Overall, content specialists were widely recognized for their publications in the area of traumatic brain injury and executive functions and had published in renowned scientific journals. A high percentage of experts judged the definitions of task and operations, the rating scale, and the instructions to be both clear and pertinent. However, experts recommended that certain modifications be made to the prototype (version 1.0). For instance, instructions for the “doing daily house cleaning” task were judged to be unclear. We therefore removed the task from the final prototype (version 2.0) and replaced it with “cleaning up after the meal,” which was integrated into the series of tasks linked to the goal of hosting a meal. Fifty-seven percent of the experts found the new six-level rating scale, modified from the original four-level scale of the ADL Profile, clear. The remaining 43% questioned the appropriateness of the proposed scale modification. More specifically, they questioned the idea of representing the need for verbal assistance as a greater state of dependence than the need for physical assistance. Based on these findings, the rating scale of the final prototype (version 2.0) was a five-level ordinal scale (dependence, verbal and physical assistance, verbal or physical assistance, independence with difficulty, and independence).

Modifications made to the final prototype to be submitted to the reliability study were done in accordance with the methodology proposed by Benson and Clark (1982), who propose that modifications be made to improve the tool under development after content validation and pilot testing.

**Reliability study results.**

Sociodemographic and clinical characteristics are depicted in Table 1. As illustrated, the average age of the subjects was 40.9 years and the majority of them were men (77%). Sixty-three percent had severe TBI, and the average time post-injury was 12.9 months. One outlier at 240 months post-injury was removed to avoid skewing the results.

Analyses of the final prototype’s (version 2.0) 29 items showed that when the ratings of R1–R3 (11) were compared to the main evaluator (R4), most of the kappa coefficients showed substantial to almost perfect agreement (2% slight, 2% fair, 6% moderate, 33% substantial, 56% almost perfect) (Table 2). Paradoxically, 100% of the kappas considered as fair (0.21 – 0.40) or slight agreement (0.00-0.20) had a percentage agreement superior to 71%. This is in line with a known behaviour of the kappas (Cicchetti, 1988). In our study it represents near perfect homogeneity of the rating. For example, though item 504 (having a meal with guests: attainment of the initial goal) had a kappa of 0.00, subjects were nearly systematically (98%) rated “independence without difficulty” by all raters. The lack of variance in the ratings, though reflecting excellent agreement between raters, does not translate to high kappas (Cicchetti).

Similar results were obtained when the ratings of R1–R3 were compared to each other as most of the kappa coefficients showed substantial to almost perfect agreement (2% slight, 2% fair, 6% moderate, 24% substantial, 56% almost perfect agreement). When the two ratings of each rater were compared to each other, 94% of kappa coefficients showed almost perfect agreement.

**Discussion**

The present study was prompted by the need to develop a new performance based-measure of IADL independence (the IADL Profile) administered in the home and community environments of adults with moderate and severe TBI and to establish its reliability. A pilot study showed the tool to be acceptable for moderate and severe TBI subjects and feasible for administration in a single session in subjects’ home and community environments. Results of the reliability study indicated that interrater agreement of the scoring varied between moderate and almost perfect. We will now attempt to situate this new instrument within the context of other existing IADL tools.
Development of the IADL Profile was based on current knowledge of TBI, IADL, and executive functions. The instrument consists of eight tasks known to reflect everyday activities in which large percentages of individuals with TBI require assistance. This includes meal preparation, housekeeping, shopping, and personal finances. When compared to previous performance-based instruments (see Fisher, 2003) and studies (see Chevignard et al., 2000), a larger number of tasks known to be relevant and problematic for individuals with TBI (see Dawson & Chipman, 1995; Kozlowski, Pollez, Thevenon, Dhellemmes, & Rousseaux, 2002) were included in the IADL Profile. As environmental barriers have been shown to limit IADL independence of individuals with TBI living in the community (Ashley, Persel, & Clark, 2001; Dawson & Chipman), the administration context selected was the individual’s home and community environments.

Table 2
Interrater Agreement - Three Raters (R1–R3) Compared to Main Rater (R4): Scale 0-4 (n=28)*

<table>
<thead>
<tr>
<th>Item Number</th>
<th>Item Description</th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Weighted kappa</td>
<td>% agreement</td>
<td>Weighted kappa</td>
</tr>
<tr>
<td>101</td>
<td>Putting on outdoor clothes</td>
<td>0.94</td>
<td>96.4</td>
<td>0.94</td>
</tr>
<tr>
<td>102</td>
<td>Formulate goal</td>
<td>1.00</td>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>103</td>
<td>Plan</td>
<td>0.85</td>
<td>82.1</td>
<td>0.89</td>
</tr>
<tr>
<td>104</td>
<td>Carry out task</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>105</td>
<td>Verify attainment of goal</td>
<td>0.86</td>
<td>85.7</td>
<td>0.83</td>
</tr>
<tr>
<td>106</td>
<td>Going to grocery store</td>
<td>0.75</td>
<td>75.0</td>
<td>0.88</td>
</tr>
<tr>
<td>107</td>
<td>Formulate goal</td>
<td>0.73</td>
<td>71.4</td>
<td>0.81</td>
</tr>
<tr>
<td>108</td>
<td>Plan</td>
<td>0.70</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>109</td>
<td>Carry out task</td>
<td>0.70</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>110</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
<tr>
<td>111</td>
<td>Shopping for groceries</td>
<td>0.77</td>
<td>82.1</td>
<td>0.77</td>
</tr>
<tr>
<td>112</td>
<td>Formulate goal</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>113</td>
<td>Plan</td>
<td>0.70</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>114</td>
<td>Carry out task</td>
<td>0.69</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>115</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
<tr>
<td>116</td>
<td>Preparing a hot meal for guests</td>
<td>0.85</td>
<td>92.9</td>
<td>0.90</td>
</tr>
<tr>
<td>117</td>
<td>Formulate goal</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>118</td>
<td>Plan</td>
<td>0.70</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>119</td>
<td>Carry out task</td>
<td>0.69</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>120</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
<tr>
<td>121</td>
<td>Having a meal with guests</td>
<td>0.77</td>
<td>82.1</td>
<td>0.77</td>
</tr>
<tr>
<td>122</td>
<td>Plan</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>123</td>
<td>Carry out task</td>
<td>0.69</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>124</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
<tr>
<td>125</td>
<td>Cleaning up after meal</td>
<td>0.77</td>
<td>82.1</td>
<td>0.77</td>
</tr>
<tr>
<td>126</td>
<td>Formulate goal</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>127</td>
<td>Plan</td>
<td>0.70</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>128</td>
<td>Carry out task</td>
<td>0.69</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>129</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
<tr>
<td>130</td>
<td>Obtaining information</td>
<td>0.85</td>
<td>92.9</td>
<td>0.90</td>
</tr>
<tr>
<td>131</td>
<td>Plan</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>132</td>
<td>Carry out task</td>
<td>0.69</td>
<td>71.4</td>
<td>0.80</td>
</tr>
<tr>
<td>133</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
<tr>
<td>134</td>
<td>Making a budget</td>
<td>0.85</td>
<td>82.1</td>
<td>0.85</td>
</tr>
<tr>
<td>135</td>
<td>Plan</td>
<td>0.85</td>
<td>82.1</td>
<td>0.85</td>
</tr>
<tr>
<td>136</td>
<td>Carry out task</td>
<td>0.73</td>
<td>71.4</td>
<td>0.73</td>
</tr>
<tr>
<td>137</td>
<td>Verify attainment of goal</td>
<td>0.67</td>
<td>75.0</td>
<td>0.71</td>
</tr>
</tbody>
</table>

* According to Landis & Koch (1977), items in bold represent moderate to almost perfect agreement beyond chance.
When developing the test, particular attention was given to including task complexity and novelty, elements known to be of fundamental importance to tests of IADL independence in consideration of executive functioning (Burgess, 2000; Crépeau, Scherzer, Belleville, & Desmarais, 1997; Goel, Grafman, Tajik, Gana, & Danto, 1997; Rabbitt, 1997). A non-structured approach (Lezak, 1989; Lezak et al., 2004) and real-world environments (Burgess et al., 2006; Burgess & Robertson, 2002) were also judged important. One particularly new contribution of the IADL Profile in relation to previous studies is the integration of all three elements into one complex and long-lasting IADL. For instance, based on a minimally structured approach, subjects were left to formulate task goals and subgoals and act on their own initiative to go to the grocery store, shop for groceries, prepare a hot meal for guests, have the meal, and clean up after the meal after having been told that they were to “get ready for guests to arrive” and that “$20 would be provided to cover incurred costs.” To document relevant goal formulating and planning behaviours, subjects were simply asked to verbalize what they were going to do. Elements of complexity were interrelated to this minimally structured approach as subjects were expected to complete all aspects of planning of combined tasks and subtasks (i.e., consider different menu options and choose meal, verify presence of necessary ingredients, decide on what to purchase, decide on where to purchase ingredients and on how to get to the grocery store). Finally, the third element, “real world environment,” was combined with the two previous elements as the test was administered in subjects’ home and community environments. This further contributed to the complexity of the IADL as subjects had to consider, for instance, the distance from their homes to the nearest grocery stores and the transportation available to reach their destination. When compared to previous studies, greater planning (identification of alternative forms of transportation for individuals living far from the grocery store without a valid driver’s license) was required for the task “going to the grocery store” when measured with the IADL Profile. This approach is more ecological, and, we expect, a closer approximation of the individual’s real-world functioning IADL independence. It is also more sensitive to the repercussions of executive functioning on the ability to accomplish this task. Also, two other tasks, “obtaining information” and “making a budget,” were designed to tap into elements of complexity and novelty although here more limited interactions were required with the real-world environment, and the test approach was more structured as goals were explicitly formulated by the examiner.

To illustrate the distinctive nature of the IADL Profile approach from other IADL assessments, let us compare the instructions of the IADL Profile to the Executive Function Performance Test (EFPT) (Baum et al., 2008), a measure that has a number of similarities with the IADL Profile but that was validated with a different patient population, that is, individuals with dementia. In the IADL Profile, instructions provided to subjects merely state that they are to prepare to receive unexpected guests. They are also informed that they will be given $20 to cover incurred expenses. With this, they need to plan not only the meal they wish to prepare for their guests, but also the ingredients they will need, where they must go to purchase these ingredients, and the means they will use to get to the selected store. The EFPT, on the other hand, gives the following more structured instructions:

I want you to make oatmeal. Here is an enlarged version of the instructions. Follow these directions and when you are done, put the oatmeal in a bowl. The items you need are in the box.

Though the goal of the test is somewhat comparable to that of the IADL Profile, the use of such instructions removes the possibility of observing the subject’s ability to formulate goals and plan, as subjects are merely required to follow step-by-step instructions and use material that is provided by the examiner. In our perspective, instructions given when administering the IADL Profile permit a more thorough evaluation of the person’s independence in everyday tasks within home and community environments.

Results of the reliability study support the need for advanced training of the raters and use of a highly detailed scoring manual. These results are partly explained by the inherent challenges presented to therapists when asked to score behaviours not traditionally documented in other tools, that is, behaviours related to goal formulation and planning. It is possible that this need for advanced training will limit the access occupational therapists will have to the tool. However, future work will be aimed at developing materials and publishing articles that may contribute to reducing the amount of training needed.

Study Limitations
Due to recruitment difficulties of moderate and severe TBI subjects, the sample sizes of both the pilot and reliability studies were small.

Conclusion
In this paper we presented the development, content validation, intra- and interrater agreement of a new performance-based measure of IADL, the IADL Profile. Based on earlier work on the ADL Profile, this new measure was developed to better estimate independence in IADL and the repercussions of executive deficits on IADL for individuals with TBI. As administration of the test is undertaken in real-world environments, we expect that it will facilitate links between assessment and rehabilitation and contribute to the development of more targeted community-based treatment interventions. This study has provided strong preliminary evidence for the IADL Profile’s reliability when administered by occupational therapists who have participated in a three-day training workshop. This study adds to previously published studies of the tool’s criterion-related and factorial validity and its generalizability, ensuring its readiness for clinical use.
Key messages
- Due to the central importance attributed to executive functions for independent living, it is essential that measures of independence in IADL capture potential repercussions of executive function deficits.
- A new tool, the IADL Profile, was developed to measure independence in IADL and the repercussions of executive function deficits.
- Scoring of the IADL Profile was shown to be reliable when administered by occupational therapists who participated in a three-day training workshop.

Acknowledgements
The authors would like to thank Mariama Touré, Emilie Lemay-Brault, Jessica McGuire, Réjean Prévost, and Dr. Julie Lamoureux for their contribution to the study. We would also like to thank all participating rehabilitation centers who recruited subjects for this study. This study was completed to meet the partial requirements of the first author's doctoral degree.

References
Bottari, C. (2001). Perception of experts on criteria for the assessment of the “dysexecutive syndrome” in everyday activities. Université de Montréal, Montreal, Quebec.
Book Review | Revue des Livres

Liz Ann Davidson, Kerrie Old, Christina Howe and Alyson Eggett
184 pages; 32.99 pounds
ISBN: 978-0-86388-583-9

Two well written workbooks provide a ready to use program for the implementation of group work with children with autism with multifaceted challenges (communication and language, socialization, play, sensory, motor, behavior and emotional). The multidisciplinary program is divided into three books in order to offer adapted information for age groups 3-5 years, 5-11 years and 11-16 years (only two were reviewed here). Each book can be used independently; however, when working with higher functioning children or those close to the starting age of the next book in the series, authors recommend to also consider activities across more than one book, thus warranting the acquisition of multiple books at once.

The program is impressively comprehensive including easy to reproduce assessment forms, sample letters, session plans, session notes, detailed activity suggestions, group evaluation forms, completion certificates and lots more. Theory is provided with regards to autism as well as to the logistics of group planning (funding, finding a location etc.). The information is presented in a logical order, thus making it easy to find and retrieve desired information very quickly. Numerous case examples are provided throughout the book to help the reader visualize applications of what they are reading. The design, layout and illustrations make these books very easy to read, photocopy and use.

In sum, these books are an essential resource for autism practitioners wishing to plan and offer group work for children in the Autism Spectrum Disorder. Furthermore, the reader will find valuable that many of the activity ideas can be easily adapted for use in individual therapy, making these books are wonderful clinical resource beyond group work.

**Josiane Caron Santha**