This is a postprint version of the following published document:


DOI: [10.1007/978-3-319-64107-2_9](https://doi.org/10.1007/978-3-319-64107-2_9)

© 2020 Springer International Publishing AG 2017
The Automated Box and Blocks Test
an Autonomous Assessment Method of Gross Manual Dexterity in Stroke Rehabilitation

Edwin Daniel Oña†, Alberto Jardón, and Carlos Balaguer
Robotics Lab, University Carlos III of Madrid,
Avda. de la Universidad, 30 (28911) Leganés, Madrid, Spain
'ona@ing.uc3m.es

Abstract. Traditional motor assessment is carried out by clinicians using standard clinical tests in order to have objectivity in the evaluation, but this manual procedure is liable to the observer subjectivity. In this article, an automatic assessment system based on the Box and Blocks Test (BBT) of manual dexterity is presented. Also, the automatic test administration and the motor performance of the user is addressed. Through cameras RGB-D the execution of the test and the patient’s movements are monitored. Based on colour segmentation, the cubes displaced by the user are detected and the traditional scoring is automatically calculated. Furthermore, a pilot trial in a hospital environment was conducted, to compare the automatic system and its effectiveness with respect to the traditional one. The results support the use of automatic assessment methods of motor functionality, which in combination with robotic rehabilitation systems, could address an autonomous and objective rehabilitation process.

Keywords: Automated, Assessment, Assistive, Manual dexterity, Rehabilitation, Stroke

1 Introduction

Since rehabilitation is a laborious process of expensive intervention, evaluating its therapeutic effectiveness is particularly important [12]. This assessment is commonly performed by health professionals themselves, using standardized tests in order to have objectivity in the evaluation, but which are susceptible to the subjectivity of the clinicians. In some cases, the evaluation methods are made up of well-defined exercises based on numerical tests (task-specific), which may be susceptible to be automated. Thus, an objective assessment of the physical condition of the subject will be obtained. Also, more time to evaluate the results will be provided to the therapist, who could correct the therapy method applied, change the level of difficulty or analyse in deep the process.

Currently, in addition to projects focused on the development of support robotic systems for rehabilitation, several research projects are also focused on automating assessment methods based on neuromotor tests commonly used in
hospitals. These traditional tests can be grouped by outcomes to be measured at any of these levels: Body functions/structure (impairment); Activities (refers to the whole person - formerly conceived as disability in the old ICIDH framework) and Participation (formerly referred to as handicap) [14]. According to this classification, some projects are developed within the Body functions category. In Otten et al. [10], an evaluation method based on low-cost sensors that record data of the user’s movements is proposed. Their outcomes were compared with those obtained using the usual Fugl-Meyer Assessment (FMA) with similar results. Their conclusions suggest that the FMA is capable of being automated. Other attempt to automate part of the FMA is presented in Wang et al. study [16], in this case it is achieved using accelerometers and focusing on scores for shoulder and elbow movements.

Other studies seek to automate tests within the Activity category. In [11] a new system is proposed, called Rejoyce Arm and Hand Function Test (RAHFT), focused on the assessment of manual dexterity. This system uses the Rejoyce work station in which the participants perform manual dexterity tasks with the help of computer games. The authors state that this is the first manual dexterity test that does not rely on human judgement, offering a quantitative, standardised and remotely manageable assessment of the results. In [5] an automatic system is implemented, that objectively assess the upper limb functionality in post stroke rehabilitation. This system performs a post-processing of experimental data obtained while performing reach tasks using the KINARM robotic system. In this case, the results are compared using the Chedoke-McMaster Stroke Assessment Scale (CMSA). In [15] an automatic evaluation based on the Wolf Motor Function Test (WMFT) is proposed. By using sensors which the users should be wearing, the time needed to complete 7 of the 17 tasks of the test is estimated. Other ongoing work is presented in [6], which aims to automate the Action Research Arm Test (ARAT). This paper has developed the automation of the ARAT’s grasp subtest, through the sensorizing of one of the objects used in the task, in this case a 7.5 cm cube. Finally, in [2] the Digital Box and Blocks Test (DBBT) is presented. It aims to automate the traditional BBT. The authors, using a Kinect® sensor, determine the number of transported blocks with a 90% success rate for 80 blocks. The system also detects the hand and its movements. However, the test administration is not addressed.

Of all the studies reviewed, several works are focused on detecting upper limb movements; either by using sensors the subject should wear; by sensorizing objects used in the test; or through computer vision systems. A common goal is to obtain automatic evaluation platforms that are objective, dynamic, that show repeatability, diagnostic capabilities and that can provide more information than the traditional scale. However, a complete automated system has not been implemented yet, and therefore, a system to be administered without any intervention of clinicians.

In this paper, the automation of the BBT scoring and the automatic administration of the test are studied. In section 2, the methodology towards an automated assessment is presented. The method to count the blocks and the
graphical interface functionality are described. The data analysis of a pilot trial results are shown in section 3. Finally, the conclusions and future works are summarized in section 4.

2 Methodology

For a rehabilitation process to be automated, the method to extract metrics and the degree of acceptance by both users and health professionals should be properly assessed. To design assistance rehabilitation systems, although the focus is on the subject to be treated, it is important to systematize the understanding of the requirements demanded by therapists in order to enable an easier integration of technology in their daily activities [13]. Regarding the method, those tests that are administered without direct contact of the professional are more susceptible to be automated. But it requires a proper patient performance, so a previous explanation of the procedure or a wizard to assist them is needed. Concerning metrics, it is essential to assess which ones give relevant information and are less invasive for the subject to be evaluated [1].

Thus, the BBT study is considered on the basis that the outcome is simple (total cubes transferred), the instructions test are systematic and clear, and the test development is well defined (three stages: training, dominant hand, and non-dominant hand). Also, for its wide use in clinical settings as an evaluation system in rehabilitation processes of people who have suffered a stroke. Complete BBT description is shown in [7].

The proposed system, named as the Automated Box and Blocks Test (ABBT) has two targets: to automate the scoring of the traditional BBT, and to enable the test administration, with the minimal participation of clinicians or without it. Thereby, two elements have been developed:

a) *Automatic test scoring*: to automate the counting of cubes, and also obtaining additional information of the subject movements.

b) *User interface*: to allow the therapist to administer the test at either the same or a remote place, to store the data obtained, to get an updated database, and to generate a historical report of previous sessions.

Based on that, the ABBT goes a step further than both the traditional test and the related study of DBBT [2], since the automatic test administration is addressed. The ABBT is made up of a portable and lightweight cube-shaped structure placed on a standard desk, as shown in Fig. 1. At the top of the structure, a Kinect® for Windows® V1 sensor is fixed. This is used for detecting the number of cubes displaced, as well as the hand movements while the subject performs the test. The BBT box is located on the desk and in the center of the structure.

The ABBT is implemented on Windows®, by using the Matlab® R2015b software. It provides, hardware support packages for devices such as Kinect® for Windows®. Besides, Matlab® software includes a set of toolboxes and libraries for image processing, and it also offers the graphical user interface development environment GUIDE to create a graphical user interface.
2.1 Automatic Test Scoring

The automatic counting process is developed in three stages:

a) Detection of box edge,
b) Segmentation by colour, and
c) Score validation.

When test starts, the algorithm looks for the edge of the box. For that purpose, the depth data of Kinect® sensor is used. As result, both the left and the right compartments of the box are identified. Based on that, a region of interest (ROI) on the RGB image is defined to be processed. According to the hand to be evaluated (dominant or non-dominant), the corresponding ROI (left or right) is selected for the cubes counting. While the test period time is not complete, the counting process is executed. When the 60 seconds are over, the results are displayed through the interface. This above sequence is used on the three stages of the test. This method, combined with voice messages allows to administer the ABBT in an automatic way.

**Detection of Box Edge** Kinect® sensor is placed on the top of the structure (see Fig.1), and there is a distance of one meter up to the desk surface. Its position is fixed by using a holder made on ABS plastic in a low-cost 3D Printer. Since the Kinect® position remains unchanged, the distance from the sensor to the border of the BBT remains constant too. The process to detect the edge of the box, and differentiate between the two compartments is shown in Fig.2.
First, depth data are captured (see Fig. 2-(a)), and after, a height threshold is applied, in this case 90 cm (distance from the box edge to the sensor). The points under this threshold are discarded, including the desk and the BBT bottom (see Fig. 2-(b)). Morphological operations to reduce noise and to label the detected areas are applied to the thresholded image. Features based on several algorithms can be extracted from the processed image. In this case, using the ‘BoundingBox’ function in Matlab, the detected rectangles and its centroids are obtained from the image. According to the centroid coordinates, it is possible to identify both the left and the right compartments (see Fig. 2-(c)). Finally, in Fig. 2-(d) the ROI on the colour image are overlapped.

The left and right compartments identification is quite important, because it is the base to define the ROI in the colour image to be processed. Also, as this ROI changes according to the evaluation of the either dominant or non-dominant hand.

Segmentation by Colour To detect and to count cubes, the captured RGB images are segmented by colour, see Fig. 3. This is based on, the high contrast existing among the cubes colours of the BBT (red, green, blue and yellow), and also between them and the bottom of the box in beech colour. The colour images are captured with a resolution of 640x480 pixels. According to the compartments identification, a colour ROI (where the BBT is placed) is cropped from the whole image. The colour ROI is considered as a matrix of size MxNx3 (see Fig. 2-(d)), and each colour is processed separately. First, the colour ROI is resized up to
three times more to improve the sample acquired. This region corresponds to the empty side of the box, and it depends on whether the test is performed for the non-dominant or dominant hand of the subject. Then, this dilated image is converted to grayscale first and to a binary map after (black and white), by tuning each colour considering the corresponding thresholds. These parameters are set automatically after the calibration of light (intensity of the scene), for which 3 cubes of each colour are deposited in one of the compartments of the box. Then a series of morphological operations (opening, closing, erosion and labelling) are applied. The same sequence (without the image capture, the cropped and the resize of image) it is repeated for each colour separately. Finally, a first count of cubes is made, depending on the colour, in the captured image. The above procedure is repeated for each captured RGB frame while the test is running.

**Score Validation** Due to the lack of sensitivity and to the high spasticity on the hands, it is possible that some individuals have trouble grabbing a single cube, and they can take two at a time. In these cases, and according to the test rules [7], the additional cubes must be discarded and must be counted as one.

Having in mind this approach, a time vector to compare very close events is used, during the performance of the test. On the basis, that a healthy individual takes about a second to move a cube, it is detected if two or more cubes have appeared in very close time instants and in periods lower than a second. In that case, the additional cubes are discarded and it is only added one to the global counter.
2.2 User Interface

Graphical interface and its elements are shown in Fig. 4. The full view of Kinect® colour sensor is presented in a small window (see Fig. 4 upper left-hand corner). The scoring of cubes obtained by each colour is also shown in the middle of the image. Furthermore, in addition to the count of the cubes by colour, there are two windows to display the results. In the first one, the cubes obtained and its times of detection for the second stage (dominant hand) are plotted (see Fig. 4 upper right corner). In the second one, and after the third test stage is completed (non-dominant hand), a comparative graph of the current session along with the previous sessions is presented (see Fig. 4 lower right corner).

In order to face the automatic test administration, the ABBT plays the test instructions through a voice synthesizer to guide the subjects during the test execution. These instructions are the same as those provided by a therapist with the traditional BBT. Respect to the patient, the interface eases the test administration. Respect to the therapist, it enables to visualize the development of cubes detection (see Fig. 4 lower left-hand corner), to offer the user’s profile options (see Fig. 4 upper left-hand corner), or to load automatically subject information stored in a local database. Among the stored information, the affected hand can be found, and therefore the ABBT select the appropriate ROI (empty region of the box).

3 Results

The ABBT system is built based on an Intel® Core i7 computer, with a Kinect® for Windows® V1 sensor. The implemented algorithm is able to process 4.22 images per second (235 ms). Counting success rate was estimated at laboratory. Since the algorithm uses the segmentation by colour for the block-counting estimation, it could lose some blocks due to changes in light conditions. After several
times of test, it is found that the system has 100% of accuracy when 30 blocks are transferred [9].

3.1 Pilot Study

A pilot study to investigate the ABTT effectiveness and impressions of individuals using the automated method was conducted at a healthcare facility. The test was administrated automatically. Healthcare professionals took part only to place the BBCT at different stages, and to count the total of cubes transferred. The manual blocks counting were compared with the total obtained by the ABTT.

The trial was conducted at Fuenlabrada University Hospital, considering the good experiences accomplished while testing another robotic assistive system in hospital environments [3]. System settings are the same as the one shown in the above Fig.1. Trials were carried out in different days in the same week. ABTT reproduces the test instructions through a voice synthesizer to guide the subjects during the test execution. These instructions are the same as the ones provided by a healthcare professional with the traditional BBCT.

As the BBCT rules shown, after the training period, the individuals proceeded to perform the test, by starting with their not affected hand (dominant). Then, the trials were carried out with their affected hand (non-dominant). At the end of each stage, one of the therapists proceeded to count the total of cubes displaced.

3.2 Participants

Four subjects with different levels of upper limb impairments were selected by medical professionals. The participants were chosen according to the following inclusion criteria: a) Affectation of the upper extremity, b) Gripping ability, c) Spasticity according to Modified Ashworth Scale \( \leq 2 \), and d) Ability to understand Mini-mental test instructions \( \geq 24 \). Demographics data of the participants in the trial are shown in Table 1.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Affectation</th>
<th>Gender</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dominant</td>
</tr>
<tr>
<td>Subject 1</td>
<td>23</td>
<td>Left-sided hemiparesis</td>
<td>Male</td>
<td>35 / 44</td>
</tr>
<tr>
<td>Subject 2</td>
<td>54</td>
<td>Akinetic-rigid syndrome</td>
<td>Female</td>
<td>45 / 56</td>
</tr>
<tr>
<td>Subject 3</td>
<td>55</td>
<td>Right-sided hemiparesis</td>
<td>Female</td>
<td>36 / 50</td>
</tr>
<tr>
<td>Subject 4</td>
<td>58</td>
<td>Right-sided hemiparesis</td>
<td>Male</td>
<td>33 / 54</td>
</tr>
</tbody>
</table>

Average of ABTT detection success rate: 73% 91%

*Scoring for the ABTT and the BBCT (bold)
3.3 Pilot Trial Results

The trial results were grouped by the relevance of the information provided to the therapist, such as: traditional scoring, additional outcome (more objectives and based on user’s performance), and clinical outcome (whose interpretation can improve the clinical evaluation). A preliminary data analysis is shown in [8].

**Traditional Scoring** The results obtained by using the ABBT have been different for the case of the dominant and the non-dominant hand. The average success ratio was of 90.75% for the non-dominant hand case, and of 74% for the dominant hand case. The scoring obtained by using the ABBT and the manual counting of blocks are also summarized in Table 1, according to each subject and in the case of both the dominant and the non-dominant hand. It can be seen that the success rate was different depending on whether the exercise is performed with the unaffected or the affected hand. This variation in the rate of success is due to the greater speed of movement with the healthy hand, and that makes difficult the detection of the cubes. Other factor is attributed to changes in light conditions, that were different during the test days.

**Additional Outcome** The total counting of blocks and the time instant in which the cubes were detected is shown in Fig.5. The results are grouped by dominant and non-dominant hand for each subject. A fairly linear trend in the displacement of cubes can be appreciated, but with different slopes. Besides, from the results obtained, it is a fact that the variation among the slopes is related to the subjects health condition. For example, a major slope is observed for the case of Subject 2, with akinetic-rigid syndrome. A minor slope is obtained for Subject 4, that suggests a major affection level with respect to the other individuals who had hemiparesis.

The linear trend, can be estimated by using the simple linear regression (SLR) method. SLR considers only one independent variable employing the relation:

\[ y = \beta_0 + \beta_1 x + \epsilon \]  

where \( \beta_0 \) is the y-intercept, \( \beta_1 \) is the slope (or regression coefficient), and \( \epsilon \) is the error term. Starting with a set of \( n \) observed values of \( x \) and \( y \) given by \((x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\). Using the SLR relation, these values form a system of linear equations. If the line is forced to start at zero, then the system could be simplified as:

\[ Y = B \cdot X \]  

where \( B \) is the slope or regression coefficient. Then, if \( X \) is the time axis, and \( Y \) is the number of cubes detected in the equation (2), by applying the SLR to the results obtained with ABBT, the relation is:

\[ N_d = V_d \cdot t \]
\[ N_{nd} = V_{nd} \cdot t \]  

(3)
where, $V_d$ and $V_{nd}$ are the slopes by the linear fit, and it represents the average speed of blocks displacement for both the dominant and the non-dominant hand, respectively.

In Fig.6, the analysis of data obtained for Subject 2 case are presented. In this case, the average speeds in transferring blocks are calculated by using equation (3). The results achieved with the BBT (see blue and yellow columns on Fig.6 right-hand side) are compared with the ABBT results.

Clinical Outcome Despite of not getting a higher rate of success on blocks estimation, the results can also be assessed in order to evaluate if the additional information provided, is relevant to the health professionals. For example, the time information of the subjects for the non-dominant hand case (see Fig.5-(b)), suggests a reaction to fatigue during the test, and is a relevant factor for therapists as an indicator of spasticity, according to the local slope changes in the time intervals (elapsed time from one block to another) of cubes transferred.
Fig. 6: Comparative between the ABBT and the BBT by its information provided for Subject 2 data.

Since, the average speed of cubes displacement for both the affected and the non-affected hand is obtained (as \( V_d \) and \( V_{nd} \) coefficients given by making a linear fit), when the value of the two coefficients are close, it could mean that the therapy will have been effective. Also, taking a look at equation (1), it can be appreciated that the \( e \) introduce a factor of non-linearity. Then, the instantaneous speeds, that depends on the time periods elapsed to transfer a block and the next one, while running the test, are not constants. It could be used as an indicator of coordination in the arm movements, related to the dispersion of the samples.

Some pictures of participants during the performance of the pilot trials are shown in Fig. 7.

4 Discussion and Conclusions

In this paper, the procedure to automate the assessment method of BBT, including the hardware and software development, is presented. Besides, the automatically administration of the test, without intervention of healthcare professionals, is evaluated. In this way, a pilot study was designed and conducted to investigate the ABTT effectiveness in a real situation. Also, to assess the clinical viability, with a minimal or null intervention of clinicians, of the automatic administration of the ABTT is intended.

In addition, the sensorization of the system is appropriate and it provides additional information respect to the traditional BBT. This information, is stored on the patient’s register directly, making easier to generate a medical history updated. In this way, the cubes colour and the time instant when they were detected are obtained and registered by the ABTT. The linearity in the timeline of cubes detection, suggest an indicator of a major level of coordination, according to the data dispersion with respect to the linear fit; and the average speeds, given by the trend slopes, can be related with the effectiveness of therapy. This
approach, requires more trials to be consolidated, but it is encouraging. Additionally, a minimum of 5 participants will be required to conduct an usability study [4].

Regarding to the DBBT [2], not only an algorithm for counting cubes was developed, but also an interface which allows to perform all stages (training, dominant hand and non-dominant hand) of the test. Through this interface the test instructions are provided to users by voice messages, and it allows to administer the test automatically. In case of failure or user distraction, due to simplicity of the interface, it allows the user to repeat the test easily. The accuracy of counting the cubes is improved too, by obtaining a success rate of 100% up to 30 cubes transferred. It is appreciated too, that the algorithm effectiveness in clinical trials has been reduced with respect to that obtained in the laboratory. This decrease is attributed on the one hand, by failure in counting all cubes due to the higher speed in the case of dominant hand, and on the other hand, by influence of ambient light conditions that it introduced false positives in the cubes detection while performing the tests. The calibration process must be improved in future developments.

After the pilot study, the interface, the instructions provided by the interface and the results were evaluated by the users and the therapists. They positively accepted the approach to automate the test, and raised suggestions and comments. Since ABBT is a low-cost system, by its few required elements, the pos-
The possibility of performing the rehabilitation assessment at home was highlighted by the participants. Their opinions and the obtained results, further supports the feasibility of the use of automated systems in assessment of stroke physical rehabilitation. This fact combined with an adequate robotic rehabilitation system, will lead to an autonomous and objective rehabilitation process.

Acknowledgments. The research leading to these results has received funding from the ROBOHEALTH-A project (DPI2013-47944-C4-1-R) funded by Spanish Ministry of Economy and Competitiveness and from the RoboCity2030-III-CM project (S2013/MIT-2748), funded by Programas de Actividades I+D en la Comunidad de Madrid and cofunded by Structural Funds of the EU.

References


