Contribution of Bayesian Inference and Single Case Research for Validating the Effects of Support Programs Intended for Children with Autism Spectrum Disorder (ASD)

Marion Wolff[1], Célia Nézereau[2], Maria Pilar Gattegno[3][5], Marie-Anna Bernard Paulais[4]
and Jean-Louis Adrien[5]

[1] University of Paris, Faculty of Fundamental and Biomedical Sciences, Ergonomics, UMR 9010 G. Borelli Institute
and ESTIA, Institute of Technology, Paris, France
[2] Psychology Office ESPAS-IDDEES, Neuilly sur Seine, France
[3] Psychology Office ESPAS-IDDEES, Bordeaux, France
[4] Psychology Office ESPAS-IDDEES, Pont-Ste-Maxence, France
[5] University of Paris, Institute of Psychology, Laboratory of Psychopathology and Health Process, LPPS, EA 4057,
Paris, France
Email: `marion.wolff@u-paris.fr`

Abstract. The aim of this paper is to present original and valid statistical methods to meet the needs of psychopathology, more particularly in order to study the effects of support programs intended for children with Autism Spectrum Disorders (ASD). Two case studies are presented, as part of a longitudinal data collection and monitoring process about children with heterogeneous profiles, each using a different support program: a CHIPPS program (Coaching Home Intervention Progress Progression School) with a large sample for the first one, and an ABA program (Applied Behavior Analysis) with a small sample for the second one. For estimating the validity of children's progress, Bayesian inference (gains analysis and prediction) has been implemented for the first study, and the Single-Case research (analysis of individual profiles) for the other.

Keywords: Autism Spectrum Disorders (ASD), support programs, effect size analysis, Bayesian inference, single-case research.

1   Introduction

Assessing the validity of progress concerning the children with Autism Spectrum Disorders (ASD) is always very difficult, because in the field of psychopathology, and more particularly of autism, the samples are often small and/or heterogeneous, due to the multiplicity of forms of ASD. Lombardo, Lai, and Baron-Cohen (2019) have noted that this heterogeneity is present at several levels of analysis such as genetics, neural systems, cognition, behavior and development, as well as in clinical features (e.g., response to treatment, outcome). Furthermore, concerning analysis of results, there is often no possibility of constituting a control group. Indeed, in terms of ethics, how can one set up a control group if, for example, the goal is to test a new method of accompaniment and learning assistance for children? It is not conceivable to ask parents for their child's participation in an experiment related to a new method of learning assistance, while stating that he/she will not be exposed to this new method because he/she will be part of a control group… Indeed, parents are often distraught by the disability of their child and are ready to test any novel treatment. For all these reasons, experimental design, data processing and results analysis present therefore recurring difficulties.

Either the samples are really too small and too heterogeneous, and in this case analyses are only descriptive (with some attempts with non-parametric tests), or they are sufficiently large and the data allow performing statistical significance tests. But the problem remains as for evaluating the children's progress in a validity and reproducibility way. It is also currently difficult to validate evaluation scales to standardize them. But using multivariate analyses and classifications (Wolff, 2003; Le Roux & Rouanet, 2004), it is possible to do so if the sample is large enough (Adrien, 2007; Cappe, Wolff; Bobet, & Adrien, 2011; Degenne, Wolff, Fiard, & Adrien, 2019).

In any event, many authors question this heterogeneity of autistic symptoms and also the inability to form control groups. Then they suggest the use of "big data" methods, which is a very attractive idea and could be very helpful for analyses (Lombardo *et al.*, 2019). But this approach requires that all people currently conducting studies on autism agree to share their data. To date, this is not yet easy to set up and it would still take a few more years for people to join. Other use evaluative method for evaluating and determining Evidence-Based Practice (EBP) in autism (Reichow, Volkmar, & Cicchetti, 2008) and different forms of validity can be used.

This paper aims at two case studies analyzed with two different statistical methods: the first one for estimating and predicting individual children's progress with "Bayesian method", and the second one for assessing the progress with the "Single-Case research". In both cases, the results were analyzed following an individualized support program for ASD children (Nézereau, 2017; Bernard-Paulais *et al.*, 2018).

The objective is not to compare support methods for children with autism, or to promote one or another method, but only to be able to propose possible solutions to analyze changes in the behavior and cognitive abilities of children with ASD who are psychologically assisted.

Without submitting mathematical demonstrations, these two specific methods will be first presented before a brief reminder on the specificity of autism and the current accompaniment programs, finally each example will be exposed.

2   Theorical Foundations of Statistical Methods

2.1 Bayesian Methods: from Significance Tests (Frequentist Approach) to Bayesian Inference

Without entering the debate (see Rouanet *et al.*, 1998; Robert, 2013), where several theories exist as to statistical inference from which stems the *frequentist approach* associated with significance tests (or inference tests or even hypothesis tests), only the main features of *Frequentist inference* vs. *Fiducial Bayesian inference* (or *Bayesian inference*) and their usefulness will be shown below.

Frequentist inference includes the tests we all know, for example among the most used ones: *Student's t-test, Anova F-test,* and associated *Confidence Intervals* (CI), *Pearson's Chi-squared test*, etc. These tests examine the difference between two groups (or more, according to the context and to appropriate tests) to determine if this difference is statistically significant. Statistical significance is established by estimating the probability of error (*p-value*) based on a ratio given by the test used. The significance level (*p-value*) is initially a proportion, in the sampling distribution of test distributions, of the random samples that are more extreme than the observed sample, from the viewpoint of the statistics being considered (mean, proportion, etc. as appropriate). In the case of a hypothesis test, this level is the probability of observing a sample as extreme as the observed sample, if the null hypothesis (H0)[1] is true. It is therefore the probability of error in rejecting the null hypothesis[2] (Lecoutre, 1998; Corroyer & Wolff, 2003).

These inferential tests can only be used provided several conditions are met: each sample, is randomly and independently selected, the data of each sample are drawn from a normally distributed population with respect to the variables studied (Gauss curve), finally we assume the homogeneity of variances (homoscedasticity)[3] and the number of individuals is sufficiently large. If these conditions are not met, non-parametric tests can still be used (among the best known ones: *Mann-Whitney's U-Test, Wilcoxon's T-Test, Kuskal-Wallis' Anova*, etc.).

Frequentist inference only indicates that the groups significantly differ from one another, and this result obtained with the tested sample can be generalizable to a larger unobserved population (which would have the same characteristics as the sample), but does not indicate anything about the magnitude of the effect (Corroyer & Wolff, 2003, Poitevineau, 2004; Rouanet, 1996; Rouanet *et al.*, 1998). An important aspect

---

[1] In the inferential statistics model, the null hypothesis is a position statement that there is no difference between two (or more) measured occurrences, or two (or more) groups, or no relationship between groups of data.

[2] The traditionally used significance level is the p =.05 (5%). The test is then significant if the observed sample is one of the most extreme 5% of samples in the test's sampling distribution. It can of course be smaller (p =.01, p=.001, p =.0001, etc.).

[3] The common test for normality is *Shapiro-Wilk's* test and for homogeneity of variance, *Levene's test.*

of statistical inference is making predictions. In the case of a difference observed between two groups (d_{obs} effect), it is essential to know what it would be possible to observe if new data were collected (Verhagen & Wagenmakers, 2014). Hypothesis testing procedures can be used to successively test different hypotheses on the unknown population parameter in order to approximate the population mean or d_{obs} effect. But using a single sample for estimating a population value is a constant issue. A CI can give a solution to this issue because it provides a range of likely values that may contain the population parameter for a given confidence level, such as the 95% level, which is the most commonly used (but others such as 90% or 99% confidence levels can also be used). Thus, a CI is defined by two confidence limits and the intervals' results would range the population parameter in approximately 95 % of the cases[4].

The frequentist solution, given by the classical hypothesis tests, is based on probabilities of sampling that range from the unknown (the searched parameter) to the known (the observable, for example d_{obs} value). Fiducial[5] Bayesian inference, proposed by Fischer (1935) offers an intuitive and natural solution, based on probabilities ranging from the known (the data, the observable) to the unknown (the searched parameter; see Rouanet, 1996 and Rouanet *et al.*, 1998 for a theoretical discussion).

Bayesian inference can be defined as an approach aiming to calculate an inverse probability without referencing prior probability distributions. The inversion of probabilities is carried out using a classical method, based on Bayes' theorem (1763). The Bayesian approach can therefore be described as "a learning model for the researcher in the face of reality" (Bernard, 1994). If the objective is to perform a prediction about a replication of an experiment, several sets of data will be considered (Rouanet *et al.*, 1998).

To better understand the difference between these two approaches, let us use the simple example of a coin flip.

If the base is a frequentist modeling, then there is a "real" probability of drawing heads, which has the value *p*. If, for example, a person draws heads six out of ten times, then the probability of obtaining heads from the results of this experiment is equal to $6/10 = 0.60$.

According to the Bayesian approach, if the coin is thrown then *a priori* probability of drawing heads is the same as that of drawing tails, i.e. $1/2 = 0.50$. This *a priori* probability is obtained from the results of other experiments carried out in the past. It is indeed obvious that the probability calculated by the frequentist method will converge towards 0.50 if the coin is thrown a large number of times.

The Bayesian method builds for each observable gain, a distribution of probabilities (Bayesian distribution) obtained by the application of the Bayes theorem. On the other hand, it is also possible to set a desired guarantee (probability) to obtain potential gains for this specific guarantee. This Bayesian distribution will probabilistically interpret what the data teach about the true value of the population gain[6].

A *Credible Interval* (CrI); equivalent of the frequentist Confidence Interval) can also be calculated.

In a frequentist Confidence Interval (CI), probability comes *before* collecting the data, i.e. there is a 95% - or 99% or…If 100 samples of data are collected and the mean value of the parameter is calculated in each, then in 95 samples, the CI contains the true mean value of the parameter in the population (value of interest: d_{obs} effect or a proportion or…). The statement about the probability is that the interval contains this true parameter value given by the observed data.

In a Bayesian Credible Interval (CrI), probability comes *after* collecting the data, i.e. is based on the data: there is a 95% - or 99% or... - probability that the true parameter value (value of interest: d_{obs} effect or a proportion or…) is located within the interval. This is more natural and more intuitive because the probability statement regarding these data comes after they have been observed (Rouanet *et al.*, 1998).

Fiducial Bayesian inference is based, from a computational viewpoint, on the results of frequentist inference, and it can be used to reinterpret many of the frequentist procedures (Casella & Berger, 1987; Rouanet, 1996; Lecoutre, 1998).

For example, if the comparison of two means from independent groups (with the usual assumptions of normality of the sampling distribution of the parameter and equality of variances) is carried out, the Bayesian CrI can be identical to the frequentist CI.

---

[4] A confidence calculated at a $1-\alpha$ level can be considered as the inverse of a significance level, α (i.e. *p-value* =.05 for 95% confidence level).

[5] From Latin *fiducia* (confidence, faith).

[6] For Bayesian calculation, a web application is available: LePAC (Lecoutre & Poitevineau, 1992): http://lmrs.math.cnrs.fr/Persopage/Lecoutre/PAC.htm.

Furthermore, in frequentist inference the significance level only indicates a statement about the sign of the difference (d_{obs} effect) and doesn't give any information about the true size (magnitude) of the effect.

Confidence intervals and Credibility intervals are not equivalent, even when they might in particular cases be arithmetically analogous (Rouanet, et al., 1998), and Jaynes (2003) indicates that confidence intervals are acceptable as inferences only in the singular cases where they are in agreement with the Bayesian intervals.

2.2   Single-Case Research

In order to complete the descriptive statistical analyses, a new statistical procedure has been implemented in order to better understand children's progress and also to give more coherence to results with a view to validation. This procedure is similar to "Single-Case Research" (SCR), which is very often used in epidemiological, educational or sociological contexts.

These disciplines share the same requirements for validation concerning a treatment, a method or an intervention but often the data can only be collected on a very limited number of subjects, sometimes even on a single subject.

This is also the case with psychopathology, where many studies cannot cover a large number of subjects, nor implement comparisons with control groups, especially to evaluate the effects of an accompanying method to improve the behavior of children with autism (Ganz, Hong, & Goodwyn, 2013).

In recent years, studies relating to the analysis of such individual protocols propose new statistical indicators which can give relevant complements to the elementary descriptions conventionally used up to now (Parker, Vannest, Davis & Sauber, 2011).

The principle is based on the measurement of "*Nonoverlapping data*" collected in two distinct phases: for example, before a treatment or application of a method (phase A), and after intervention (phase B). If there has been real progress, the values obtained must be distinguishable (for example, a subject that would not progress would obtain the same grade in phases A and B) and the size of the effect (the mean effect, the difference between A and B) must be higher if the treatment or method has been effective.

Most of the time, these differences are commented on the basis of previous experiences (the so-called *semantic criterion*) or in relation to the results of other subjects or in using other indicator, such as the *scaled effect* (*Cohen's d*): d/s (where *d* represents the difference between 2 mean scores and *s* the scores' dispersion into the groups). Cohen (1988, 1992) indicates that an effect around 0.20 can be considered as a weak effect, and an effect beyond 0.80 as a strong effect. It is also possible to "translate" Cohen's proposal by suggesting limit values: an effect will be considered negligible from 0 to 0.35, intermediate between 0.35 and 0.65 and strong beyond 0.65 (Corroyer & Wolff, 2003).

The nonoverlap principle has already been used by Scruggs, Mastropieri, & Casto (1987) who proposed the *Percentage of Nonoverlapping Data* (PND) index. PND is a very widespread index and very easy to implement (in phase B, the data superior to those obtained in phase A are enumerated then reported to the total workforce). According to Scruggs and Mastropieri (1994), if PND > 70%, the effect can be considered important and if PND < 50%, the effect will be negligible. However, PND does not allow inference (Parker, Vannest, & Davis, 2011).

To overcome this limitation, Parker, Vannest, Davis, and Sauber (2011) and Huskens, Verschuur, Gillesen, Didden, & Barakova, (2012) proposed a new method related to non-parametric techniques to appreciate the size of an effect (Effect Size: ES): *Tau-U*, which represents an family index that can combine the analysis of the AB phase (with nonoverlapping data) and the trend for phase B. This allows the undesirable positive trend of phase A to be controlled. The *Tau-U method* is presented as an alternative to both regression-based and nonoverlapping models because the Tau-U method also measures nonoverlap data between two phases using *Kendall rank correlation coefficient* (Kendall's Tau coefficient (Kendall & Gibbons, 1999) for matched data: Tau$_{novlap}$.

Tau$_{novlap}$ can be interpreted as the *Spearman Rhô* or the *Bravais-Pearson correlation coefficient*, where the absolute value (r = |1|) indicates a perfect correlation between the data. It can also be defined as a derivation of *Kendall's Tau* and the *Mann-Whitney U-test* (search for differences between groups of data),

since it also refers to the sampling distribution "S". The calculation of a confidence interval and of p-values are possible (precision 91-95%)[7].

The both case studies presented now after were conducted on ASD children. The first one will use the "Bayesian inference" and the second will use the "Single-Case method".

Nevertheless, a brief reminder of the characteristics of autism, and the current accompaniment's approaches, will present, in order to note the difficulty of making samples regarding to the diversity of ASD, and the challenge of implementing appropriate statistics to assess children's progress and to deduce/predict probable results for the future, using data that can never meet strict experimental criteria, such as those classically recommended in the experimental research field.

3   Autism Spectrum Disorders (ASD), Diagnosis Tools, and Accompaniment Programs

3.1   Characteristics of ASD and Diagnosis Tools

For several years, the terms "Autism", "Autistic Disorder", "Asperger's Syndrome" or "Pervasive Developmental Disorders" (PDD) were used. However, in the new classification (Edition 5th of the Diagnostic and Statistical Manual of Mental Disorders: **DSM-5**; APA, 2013), Autism Spectrum Disorder (ASD) included all these syndromes and was now defined as a "neurodevelopmental disorder".

In the American Psychiatric Association classification (APA, 2013), these names have been replaced by the notion of "Autism Spectrum Disorder (ASD)".

Since the 40's (Kanner, 1943), the definition of autism has evolved, as have diagnosis tools.

Thus, four of the five disorders indicated by the precedent version, **DSM-IV-TR**, (autism, Asperger's syndrome, childhood disintegrative disorder, and Pervasive Developmental Disorders Not Otherwise Specified: PDD-NOS; APA, 2000) are currently grouped under the general name "ASD", the levels of autism must be determined by the clinician. The second classification is the most frequently used and recommended by the High Authority for Health (HAS, 2010) is the 10th edition of the International Classification of Diseases: **ICD-10** (OMS, 1999), which includes eight types of disorders. The gender ratio is relatively constant (approximately 4 to 5 boys for every 1 girl).

Children with ASD are characterized by disorders of social communication (1st diagnostic criterion) but also by repetitive and stereotyped activities and resistance to change (2nd diagnostic criterion). This second characteristic reflects their difficulties in integrating new and changing information that destabilizes their emotional, cognitive and communicative schemas, or disorder of regulation and flexibility (Adrien *et al.*, 1995, 2001; Ozonoff, 1997).

To assess the behavior of children with ASD, a lot of quantitative scales exist. Hereafter, some of the most used in France for measuring (the list is not comprehensive):

Assessment of intellectual abilities:
- Social and Cognitive Evaluation: **SCEB Battery** (Adrien, 2007; Adrien *et al.*, 2016; Thiébaut, Adrien, Blanc, & Barthélémy., 2010, Bernard-Paulais *et al.*, 2019). This battery consists of 16 scales divided into two main sectors: 7 for cognitive functioning and 9 for socio-emotional development. Behaviors are observed for each of the 16 scales with a rating ranging from 0 to 2 (0 = failure, 1 = emergence and 2 = success). It is intended for children whose developmental age is between 4 months and 24 months.
- Wechsler Intelligence Scale for Children, 5th edition: **WISC-V** (Wechsler, 2014). This scale is used not only as an intelligence test (M = 100 and SD = 15), but as a clinical tool, for example to diagnose attention-deficit hyperactivity disorder (ADHD) and learning disabilities. It is intended for children aged 6 to 16 years 11 months
- Differential scale of Intellectual Efficiency – revised edition: **EDEI-R** (Perron-Borelli, 1996): Seven tests are proposed and the results allow the calculation of raw scores which are then converted into Efficiency Levels (EL) and Developmental Age (DA) in terms of global, verbal, nonverbal and

---

[7] For these analyses, Tau[novlap] can be obtained via a web application: WINPepi (Vannest, Parker, & Gonen, 2011): http://www.brixtonhealth.com/pepi4windows.html.

categorical abilities. ELs are standardized scores (mean $= 100$, SD $= 15$) comparable to Wechsler's Intellectual Quotients. It is intended for children aged 4 to 9 years.

- Psycho-Educational Profile 3rd edition: **PEP-3**: This tool was developed in North Carolina by Schopler, Lansing, and Reichler (2010) as part of the development of the TEACCH program (Treatment and Education of Autistic and related Communication handicapped CHildren), described in following paragraphs. It allows the evaluation oh 6 domains of development. For children aged 2 to 7 years.

Assessment of behavioral autistic symptomatology:

- For assessing the degree of severity of autism, the Childhood Autism Rating Scale: **CARS** (Schopler, Reichler & Renner, 1988) is commonly used. This scale evaluates 15 domains and determines autism severity, according to boundary scores (< 30: non-autistic; 30-36.5: mild/moderate; 37-60: severe).

Evolution of the dysregulation:

- Cognitive and emotional activities regulation assessed by Regulation Disorders Evaluation Grid: **RDEG** (Adrien,1996; Martineau *et al.*, 1992, 1998)**.** Scoring of the items between 0 – behavior never observed - and 4 (behavior very severely expressed). The total score is then categorized: score < 15: no regulation disorder; $16 < $ Score < 30: mild dysregulation; $31 <$ Score < 45: moderate dysregulation.

Evolution of repetitive and restricted behaviors:

- Repeated and Restricted Behaviors in Autism: **RRB scale**. Currently in France, there is only one validated tool for evaluating the intensity of repeated and restricted behaviors or resistance to change for patients with ASD. Created by Bourreau, Gomot, Roux and Barthélémy (2009), this tool consists of 35 items on a Likert scale ranging from 0 to 4 (0 corresponding to never observed behaviors and 4: characteristic behaviors and troubles expressed very severely). A total score is calculated, the maximum being 140.

The difficulties of ASD children appear particularly during cognitive learning, school and acquisition of autonomy, which require an integration of new elements. This disorder of regulation and flexibility activity can be reduced by expert coaching and a structured environment. Indeed, the performances of ASD children are better if an expert adult supports their activities by the repetition of the help action, the simplification of the instruction, as for example during the functional and symbolic plays, Blanc *et al.,* 2000) and if the work environment is automatically controlled (Ferster & DeMyer, 1961).

3.2 Current Accompaniment Programs

In the field of autism, the latest interventions recommended by the High Authority for Health (HAS, 2010, 2012) among people with Autism Spectrum Disorder (ASD) are part of a behavioral and developmental approach, such as the:

- **ABA** approach (Applied Behavior Analysis), which is developed on the basis of Skinner's "operant conditioning" (manipulating the consequences of a behavior, especially with the use of reinforcement) and Pavlov's "conditioning responder" (manipulating pre-behavioral stimuli), the ABA is based on principles of learning and can be used in all situations and with all populations (Leaf, & McEachin, 2006; Leaf, McEachin, Taubman, & Biesse., 2010).
- **TEACCH** program (Treatment and Education of Autistic and related Communication handicapped CHildren, Schopler, Reichler, & Lansing, 2002; Mesibov, Shea, & Schopler, 2005). The authors of the program highlight the benefits of structured and individualized learning situations for people with autism. They are specifically interested in the individual with autism, take into account their specificities and develop a program around their skills, interests and needs. More specifically, this method is based on three important aspects: early diagnosis and assessment, collaboration, and formal education.
- **ESDM** (Early Start Denver Model; Rogers & Dawson, 2013), is a behavioral and developmental intervention for ASD children aged of 12-48 months. It is based on the methods of applied behavior analysis (ABA) and on developmental and psychopathological theories related to autism.
- **CHIPPS** program (Coaching Home Intervention Progress Profession School), which offers individualized support in an ordinary environment, with the objective of promoting quality interventions for ASD people. It is an intervention program based on the models of psychoeducation (Gattegno, Fernier, Granier-Deferre, & Adrien, 2005). It is performed by psychologists who are

supervised by experienced ESPAS Psychology Service [8]. This developmental and integrative intervention program is fully in line with the recommendations of the High Authority for Health concerning "Coordinated educational and therapeutic interventions for children and adolescents" (HAS, 2012), and corresponds to all characteristics of the Naturalistic Developmental Behavioral interventions (NDBIs) which are empirically validated treatments for ASD such as argued and demonstrated by Schreibman *et al.* (2015).

For many years, the CHIPPS program has been the subject of several studies evaluating in particular its benefits with regard to the professional insertion of adults with autism (Gattegno, 2004; Wolff, Gattegno, & Adrien, 2008; Gattegno, Wolff, Ragonnet, Pascaud, de Fenoyl, & Adrien, 2017), but also the evolution of children with autism receiving support at school and at home (Gattegno, Wolff, & Adrien, 2012). Very recently, this support system has been the subject of a research on the evolution of children with ASD (Nézereau, 2017; Nézereau, Wolff, Gattegno, & Adrien, 2018).

In addition, the recent development of New Information and Communication Technologies (NICT) and more specifically of touchscreen tablets, has allowed a considerable increase in accessibility to the most recent tools for people as well typical development than for persons with disabilities. Thus, there is an increasing use of NICTs in the educational practices offered to ASD people, and their benefits are indisputable. For example, different authors are interested in the game and show that children with ASD prefer to use the tablet to play rather than tangible toys (Munoz, Barcelos, Noel, & Kreisel, 2012; Kamaruzaman, Nor, & Azahari, 2016).

Other studies highlight the fact that children with ASD learn better when they can work from technological tools such as tablets, applications, videos (Clark & Green, 2004; Delano, 2007). Most of these tools use visual aids and cues that people with autism tend to be very sensitive to, for example, verbal guidance (Bölte, Golan, Goodwin, & Zwaigenbaum, 2010). Another research conducted with children and adolescents with ASD shows that the tablet is a better learning medium for children's participation (Wolff, Gattegno, Adrien, Gabeau, & Isnard, 2014). Finally, the contribution of tablets is also emphasized by Sani-Bozkurt, Vuran, and Akbulut (2017) through the design of interactive social stories that are then presented on tablet to children with ASD. The results illustrate the acquisition of useful skills not initially targeted, as well as the development of target behaviors. Interactive stories have been a source of great motivation for children and have fostered greater autonomy.

It is the reason why the number of mobile applications available on tablets and specifically designed for ASD people continues to grow. Among all these tools, *LearnEnjoy applications[9]* are the few software programs conceptualized on the basis of theoretical and scientific foundations (Bourgueil, Regnault, & Moutier, 2015). These applications can be downloaded on a tablet (Ipad type).

LearnEnjoy applications have recently been proposed as a complementary remediation tool within the CHIPPS program, not to replace the individualized support device but to complete it (Nézereau *et al.*, 2018; Nézereau, Wolff, Gattegno, & Adrien, 2020). One of the benefits of LearnEnjoy applications is that they integrate different strategies and contexts that are recognized to facilitate the learning of ASD people, such as the need for repetition (Taubman *et al.*, 2001), a refined design (Bertone, Mottron, Jelenic, & Faubert, 2005) and the need to propose simple instructions (Murphy, 2006). If provided by an expert accompaniment, then a reduction of disorders for regulation and resistance to change can be expected (Nézereau *et al.*, 2018).

For the first case study, the children were accompanied with the CHIPPS program using the LearnEnjoy software applications via a touch pad, and for the second with the ABA approach.

4   Case Study 1: Psychological and Behavioral Evolution of ASD Children Following a CHIPPS Program Including Learnenjoy Software Applications

Over a 2-year period, this study aims to analyze the evolution of the adaptation and regulation modalities of children with ASD when using LearnEnjoy applications, which are integrated as a complementary tool to a developmental and behavioral intervention program, the CHIPPS program. The aim is not of studying

---

[8] http://www.espasiddees.fr/

[9] http://www.learnenjoy.com/

the effectiveness of these applications in comparison with others or in with none other support (Nézereau, 2017).

The hypothesis of this study is that children and adults, after an adaptation period, would show better abilities to regulate their activity and improve their cognitive flexibility using these applications in the CHIPPS program context (Nézereau, 2017; Nézereau *et al.*, 2016; 2018).

In addition, one of the objectives of this research is also to make professionals aware of the importance of evaluating the methods of care and the necessary and constant questioning of the methods of work aiming at a lasting improvement of practices.

To verify this hypothesis, test batteries from the psychopathology community were used; and to estimate accurately the progress of ASD children over a two-year duration, a specific statistical method has been implemented: Bayesian inference. The tests and the statistical method will be described in the following paragraphs.

4.1 Method

4.1.1 Participants

Fifty-one participants, including 43 boys and 8 girls[10], participated in this study. In order to benefit from the applications, these participants had to have a Global Developmental Age (GDA) evaluated using appropriate tests: SCEB (Adrien, 2007; Thiébaut *et al.*, 2010; Adrien *et al.*, 2016; Bernard Paulais *et al.*, 2019), PEP-3 (Schopler *et al.*, 2010), EDEI-R (Perron-Borelli, 1996), WISC-V (Wechsler, 2014), between 2 years and 6 years. Their chronological ages ranged from 1 year 10 months to 9 years 3 months at the beginning of the study ($M = 4.7$; $SD = 1.8$).

All subjects in the study were diagnosed with PDD/ASD according to the ICD-10 (OMS, 1999), DSM IV-TR (APA, 2000) and confirmed as subjects with ASD based on DSM-5 criteria (APA, 2013), and assessed with the Childhood Autism Rating Scale criteria (CARS; Schopler *et al.,* 1988). The diagnoses of ASD with or without ID were given by child psychiatrists and psychologists experienced in ASD and other neurodevelopmental disorders and blind to this study.

Among these 51 children, 22 have a mild/moderate ASD and 29 a severe intellectual disability (ID) - including 17 children without ID, 24 with mild/moderate and 10 with severe/profound ID as a comorbidity.

All the participants were educated in ordinary educational environment and received support through the CHIPPS program (Gattegno *et al.*, 2005; HAS, 2012), which varied from 17/20 to 35 hours per week, and they are used to working with touch pads. Some of them also benefitted from rehabilitation focused on language and / or motor skills.

4.1.2 Procedure and LearnEnjoy applications

LearnEnjoy applications come in the form of 3 modules of different levels: *Basics, Progress and Preschool*. Each of the applications offers numerous activities spread over 5 domains ("Understand and Organize", "Communication", "Live Daily", "Play and Interact", "Consider the School") and which are hierarchical in progression of lessons and their difficulties. This progressivity concerns both language and visual performance (for example in matching, sorting, categorization activities), motor skills, oral-facial praxis, imitation, etc. (Bourgueil *et al.*, 2015). See example figure 1 here after.

In order to obtain comparable data, a common course training was established to all participants. Each accompanied person follows at his own pace the same program which is composed of 10 activities among 5 domains. Thus, some subjects will do a program in one week while others will do it in two or three. The choice of the number of activities by domain was made according to the specificities related to autism (areas to be strengthened in priority). The applications offer flexibility that also allows the caregiver working directly with the person to adapt the program based on his observations in real time (frustration, weariness, failure, etc.).

Each psychologist completes the questionnaires provided by LearnEnjoy to determine which application-level (Basics, Progress and Preschool) the accompanied person should start. After a first trial, the questionnaires were refined and items more specific to autism and intermediate levels were added. Then

---

[10] As Gender-ratio

each accompanied person starts to work with the applications by following the basic program adapted to his pace and skill level.

Figure 1. Preschool level 1. Sorting categories: classify images (to the left)/ Performing praxies (to the right): go around the mouth with your tongue.

Each of the 51 children participated in 3 intellectual and behavioral assessments. Each of them corresponds, for the analyses, to the "Period" factor (from P0 to P2):

P0: First evaluation, preliminary to the implementation of the support system (before CHIPPS program).

P1: Second evaluation (after about 10 months of CHIPPS program).

P2: Third evaluation (after about 20 months of CHIPPS program).

These assessments were conducted by qualified psychologists, and the ethical principles relating to the Code of Ethics have been respected (Article 44 of French law n°85-772). They consist of:

- An intellectual ability assessment with the measure of the Global Developmental Age (GDA) and a more precise analysis of verbal skills (Verbal Developmental Age: VDA). Developmental Quotients were also calculated (Global: GDQ and Verbal: VDQ). These results are obtained using developmental assessment tools such as SCEB, PEP-3, EDEI-R, WISC scales and are expressed in months.
- A behavioral symptomatology assessment using the Childhood Autism Rating Scale: CARS. Results are expressed in decreasing scores according to the severity autism levels (< 30: non-autistic; 30-36.5: mild/moderate; 37-60: severe).

These 5 dependent variables (GDA, VDA, GDQ, VDQ and CARS) were studied according to different factors (Independent variables) as: gender, intervention site (halt-daycare, preschool education, and school), duration of support with the CHIPPS program (17h/25h or 35h per week), intellectual disability (without, mild/moderate/severe), ASD severity, and assessment period (P0, P1, and P2 as described before).

The variables GDA, VDA are Global and Verbal Developmental Age indicators from DSM-5 specifications, and GDQ or VDQ represent Quotients (for the same measures) from DSM-IV-TR specifications. These 4 variables are highly and significantly correlated, regardless of the test period (between $r(49) = 0.63$ and $r(49) = 0.93$) and results are therefore redundant. For this reason, only the results relating to GDA and VDA will be presented hereafter.

After examining the data regarding the characteristics of the children and their environment (analysis of the effects for: gender, ID, autism severity, location of intervention, and time of weekly accompaniment), the main hypothesis concerns the effect of the CHIPPS support associated with LearnEnjoy applications.

It is expected that all children would improve their results concerning their intellectual abilities, whatever the severity level of ASD or ID (best scores for GDA, VDA), and would decrease autistic symptoms (lower scores for CARS) between P0 and P2.

4.2 Results

First, after verifying normality of the data and homogeneity of variances, analyses of variance (ANOVA *F*-test) and post-hoc tests (*Scheffé* test, Student's *t*-test) were carried out, even if some values could be considered atypical, taking into account the diversity of individual profiles, but an Anova is not sensitive to moderate deviations and it is the case in this study. For estimating/predicting gains at different scales, and thus children's progress, between P0 and P2, Bayesian inference was implemented.

4.2.1 Effects of gender, intellectual disability and Autism severity

- Gender. It was verified that gender has no influence on results, which do not differ significantly between the 43 boys and the 8 girls, whatever the examined variable. (all *p values* >.10).
- Intellectual Disability. As expected, the GDAs, VDAs and CARS scores obtained by the 3 groups differ significantly [F (2, 48) = 15.57, p =.000; F (2, 48) = 20.13; p =.000; F (2, 48) = 6.32; p =.004]. The children without ID (n = 17) obtain best scores than the other two groups (n = 24 mild/moderate; n = 10 severe/profound). However, *Scheffé* post-hoc tests do not indicate significant difference between these 2 last groups. Thereafter, the analyses will therefore only be carried out with 2 groups: "without ID" and "with ID" (without distinction between the levels of developmental delay). The different descriptive indicators will be presented with comments of figure 2 and table 1.
- Autism severity. Predictably, the analysis of the variance shows a significant difference between the 2 groups of participants with mild/moderate autism (n = 22) and severe autism (n = 29) for scores relative to GDAs [F (1, 48) = 6.23; p =.01], VDAs (F (1, 48) = 11.29; p =.002] and CARS [F (1, 48) = 41.97; p =.000]. It should be noted that ID takes into account more than one-half of the data dispersion, which is not negligible (η^2 =.52). The differences will be commented here after (figures 3 and 4, table 2).

4.2.2 Effects of the environment

- Location of intervention (daycare center: n = 8; preschool education: n = 35; and school: n = 8). The location of intervention (closely related to the age of the children) indicates increasing developmental differences (from daycare center to school) for GDAs [F (2, 48) = 6.95; p =.002], which is consistent.
  However, for VDAs and VDQs, the results do not differ significantly, which is explained by the fact that the primary symptoms of autism are expressed by an alteration in language skills. The same is true for CARS scores because the symptomatology of autism does not depend on the age of the children.
- Duration of the CHIPPS program: 17h/20h per week (n = 32) vs 35h per week (n = 19). The duration has no significant effect, contrary to a legitimate expectation, although at the descriptive level it is observed that the 32 children who have received 35 hours of intervention/week tend to obtain higher GDA months-scores and lower CARS scores. This lack of a significant effect can be explained statistically by an inter-individual dispersion due to the great diversity of profiles of ASD children, where ID has a large part because the group with the longest duration of the CHIPPS support is the one that includes children with severe ID.

4.2.3 Effects of the CHIPPS program over a 2-year period (3 assessments from P0 to P2, every 10 months)

Whatever the studied variable, and whatever their profile, the children progress significantly between P0 and P2 [GDA: F (2, 100) = 106.81; p =.000 – VDA: F (2, 100) = 62.15; p =.000] and their CARS scores tend also to decrease [F (2, 100) = 68.02; p=.000]. The effect of the CHIPPS support is therefore convincing, but how estimating the children progress according to their severity levels of intellectual disability or autism? On the one hand, calculations of interaction effects will be necessary to assess within-group effects and on the other hand, Bayesian inference will provide probabilities on the reproducibility of gains for future probable data.

4.2.4 Estimating gains and interaction effects between period and intellectual disability for children following CHIPPS program

The increases are not identical depending on the child's level of disability and interaction effects with period factor are significant for ID levels (without ID/mild to severe ID), for GDA or VDA scores (in months). Table 1 of within-group effects and figure 2 indicate that, after 2 years, the children without ID

progress at twice the rate for GDA (gain = 28.5 months) of those with ID (gain = 14.2 months) and for VDA, almost three times more (30.1 *vs* 11.4 months). It should be noted that their progress is more important during the second year.

For CARS, the effect is not significant (see table 1): the children with a delay progress more after 2 years (-6.35 points) than those without an Intellectual Disability (-5.85 points) but the difference is negligible. Intellectual Disability does not appear to influence the progression of autistic symptomatology and the CHIPPS program allows all children to significantly decrease their autistic behaviors, as previously indicated, regardless of the intensity of the developmental delay at the beginning of the study. Unlike intellectual abilities whose evolution is influenced by severity' degree of ID.

Note: For the following graphics (Figures 2, 4 and 5), the red line segments ("error bars") represent the *p-value*.95 confidence intervals around mean points.
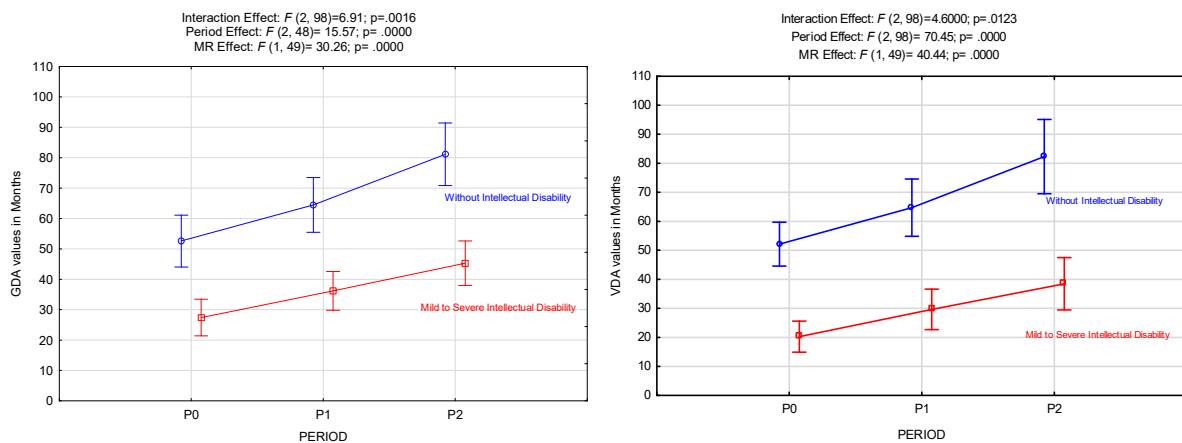

Figure 2. Interaction effects between Period (from P0 to P2) and Intellectual Disability Levels for GDA (on the left) and VDA months-scores (on the right).

Table 1. Within-groups effects of Intellectual Disability for the 3 periods (GDA and VDA months-scores, and CARS inverse-scores).

		P0	P1	P0-P1 Gain	P2	P1-P2 Gain	P0-P2 Gain
GDA	Without Intellectual Disability	52.59	64.47	11.88	81.12	16.65	28.53
Scores	Mild/Moderate Intellectual Disability	24.70	31.30	6.60	38.90	7.60	14.20
VDA	Without Intellectual Disability	52.12	64.71	12.59	87.29	17.58	30.17
Scores	Mild/Moderate Intellectual Disability	20.06	27.40	7.34	31.50	4.10	11.44
CARS	Without Intellectual Disability	37.59	34.47	-3.12	31.68	-2.79	-5.85
Scores	Mild/Moderate Intellectual Disability	46.20	42.30	-3.90	39.85	-2.45	-6.35

These results are quite consistent and indicate that coaching has a beneficial effect on learning. But what can be anticipated for the future?

In accordance with Table 1, we will only focus on progress for each group, and not on intergroup results, for the following results concerning Bayesian approach about effect size between P0 and P2 for GDAs and VDAs scores. In other terms, what are the probabilities of gains to be expected with the CHIPPS program?

Hereafter, we detail predicting GDA gains for children without intellectual disability. For the other cases, results will be summarized in tables 2 and 3. From paired samples *t-tests* and d_{obs} *effect* between P0 and P2, fiducial Bayesian inference gives the results here after (Figure 3):

It is noted that a gain of approximately 24 months may occur in 90% of cases, a gain of 23 months with a 95% probability and a gain of 20 months with the 99% guarantee (the value decreases with increasing probability), while the guarantee will still be 50% to find *d effect*, because the distribution is centered on

this effect. Therefore, the gain of *d effect* is predictable in 1 out of 2 cases for all groups of participants! If we wish to be able to predict for as many people as possible, without being too ambitious and detached from the usual level of significance, the 90% guarantee seems to be the best compromise.
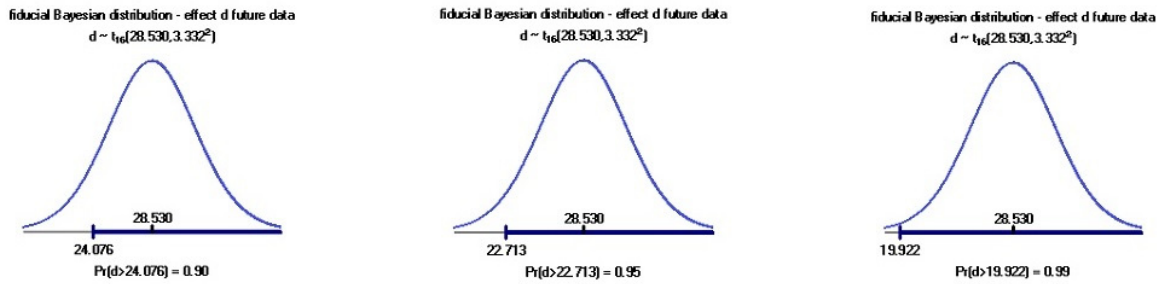

Figure 3. Predicting GDA gains for children without intellectual disability (90%, 95%, and 99% probabilities calculated with PAC Software).

In Table 2 hereafter, for each child group (without ID – or with mild to severe ID), are indicated: number of individuals (*N*), frequentist results as the *t-test* value, degrees of freedom of the test (*df*), the *p-value* of the test, observed difference between P0 and P2 (d_{obs} *effect*)[11], Confidence Interval (*CI* at 95% level), and Bayesian probabilities at 90% and 95% for the *d* effect: *Pr (d > value)* = 90% or 95%, and Credible Interval (*CrI*) at 95% level for a comparison with CI and first authorized p-value for a significant frequentist result.

For all the cases, the Credible Interval (CrI) is most of the time larger than Confidence Interval (CI), except for GDA months-scores for children with ID where CrI and CI are approximately equivalent. The gains are all notable, although there are better for the children without ID: almost a 2-year progress in GDA and VDA and a 1-year progress for children with ID! These results can be predicted in 90% of the cases. See Table 2.

Table 2. Frequentist and Bayesian results for the 3 periods (GDA and VDA months-scores, and CARS (inverse-scores) according the intellectual disability (ID) severity.

Statistical index variable	N	t-test	df	p-value	d_{obs} effect	CI 95%	CrI 95%	Pr (d>x) 90%	Pr (d>x) 95%
GDA Without ID	17	12.11	16	.0000	28.53	[23.53; 33.52]	[21.47; 35.59]	24.08	22.71
GDA With ID	34	7.66	33	.0000	17.88	[13.13; 22.63]	[13.63; 22.13]	15.15	14.35
VDA Without ID	17	6.13	16	.0000	30.17	[19.74; 40.62]	[15.49; 44.87]	20.90	18.07
VDA With ID	34	6.36	33	.0000	16.70	[12.39; 24.04]	[9.98; 26.46]	12.92	11.36
CARS Without ID	17	5.27	16	.0000	-5.91	[-8.29; -3.53]	[-9.27; -2.55]	-3.79	-3.14
CARS With ID	34	5.09	33	.0000	-7.06	[-9.88; -4.24]	[-11.05; -3.07]	-4.50	-3.74

For the CARS, which determines autism severity, it can be predicted in 90% of the cases an approximately 4-point decrease for the two groups, which is not spectacular, but nevertheless an appreciable result.

The CHIPPS program associated with the LearnEnjoy applications is therefore considered effective for children with ASD, without or with MR. The same approach is now applied to GDAs, VDAs and CARS depending on the level of severity of autism.

---

[11] Calculation direction: P2-P0.

4.2.5 Estimating gains and interaction effects between period and Autism level for children following the CHIPPS program

In this study, 22 participants had mild to moderate autistic disorder (CARS scores between 30 and 36.5) and 29 had severe autism (CARS scores above 36.5). Figure 4 here after indicates that for GDA scores (to the left) the interaction effect is not significant, but Period and Autism Level effects are significant. All children are regularly increasing their scores whatever their disorder, but for VDA scores (to the right), there is an influence of autism severity: if autism is more severe at the beginning of the study, the verbal developmental ages tend to progress much less and the gap is widening between the 2 groups of children in P2 and the interaction effect is significant. For CARS scores, children with severe autism at P0 progress more than those with mild/moderate autism (the interaction effect is also significant). These results are in concordance with those obtained in terms of Intellectual Disability: it seems consistent that children with highly intense autistic symptomatology have a greater margin of progression than others.
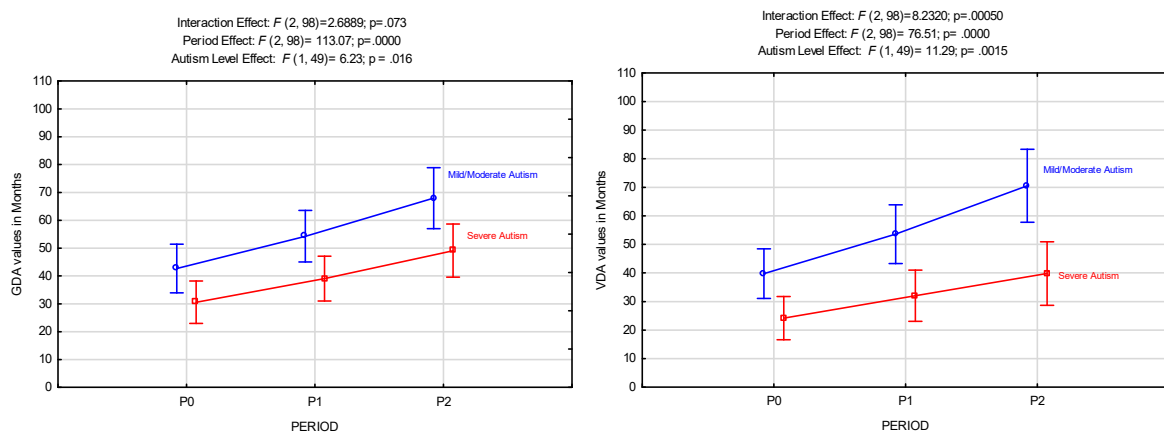

Figure 4. Interaction effects between Period (from P0 to P2) and Autism Levels for GDAs (a) and VDAs.
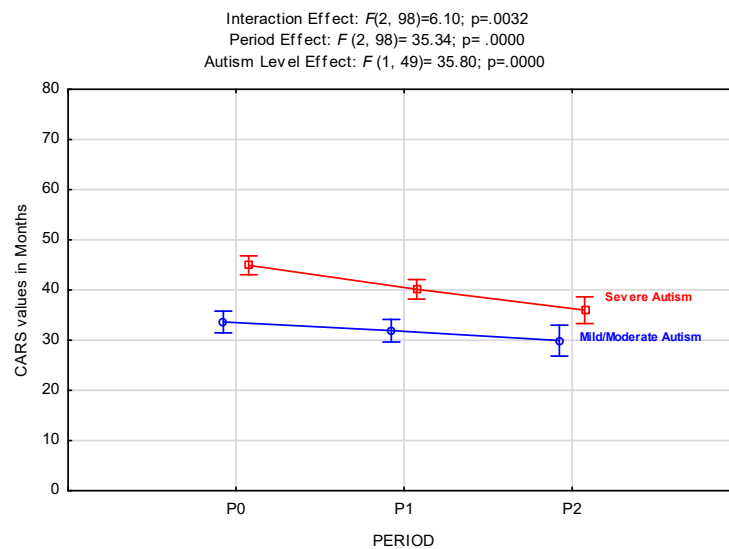

Figure 5. Interaction effects between Period (from P0 to P2) and Autism Levels for CARS.

For GDAs, within-group effects (Table 3) show a greater increase for children with mild to moderate autism at P0, since they gain about 25 months in 2 years, while participants with severe autism gain

approximately 19 months. For VDAs, children with milder autistic symptomatology double their progression compared to those with more severe disorders and the opposite is true for CARS scores.

The table 4 below indicates detailed predictions of GDA, VDA and CARS gains for ASD children. As seen in the previous paragraph, from paired samples *t-tests* and d_{obs} *effect* between P0 and P2, fiducial Bayesian inference gives the following results:

Table 3. Within-groups effects of Autism severity for the 3 periods (GDA and VDA months-scores, and CARS inverse-scores).

		P0	P1	P0-P1 Gain	P2	P1-P2 Gain	P0-P2 Gain
GDA Scores	Mild/Moderate Autism	42.68	54.27	11.59	67.95	13.68	25.27
	Severe Autism	30.59	39.03	8.44	49.1	10.07	18.51
VDA Scores	Mild/Moderate Autism	39.73	53.59	13.86	70.55	16.96	30.82
	Severe Autism	24.16	32.03	7.87	39.83	7.80	15.67
CARS Scores	Mild/Moderate Autism	33.64	31.91	-1.73	29.93	-1.98	-3.71
	Severe Autism	44.93	40.14	-4.79	37.24	-2.90	-7.69

Table 4. Frequentist and Bayesian results for the 3 periods (GDA and VDA months-scores, and CARS (inverse-scores) according the Autism Level.

Statistical index Variable	N	t-test	df	p-value	d_{obs} effect	CI 95%	CrI 95%	Pr (d>x) 90%	Pr (d>x) 95%
GDA Mild/moder. autism	22	8.83	21	.0000	25.27	[19.32; 31.23]	[16.85; 33.69]	19.92	18.31
GDA Severe autism	29	7.82	28	.0000	18.52	[13.66; 23.37]	[11.66; 25.38]	14.12	12.82
VDA Mild/moder. autism	22	6.75	21	.0000	30.82	[21.32; 40.32]	[17.39; 44.25]	22.28	19.71
VDA Severe autism	29	6.37	28	.0000	15.67	[10.63; 20.71]	[8.56; 22.79]	11.11	9.76
CARS Mild/moder. autism	22	5.20	21	.0000	-3.70	[-5.19; -2.22]	[-5.79; -1.61]	-2.37	-1.97
CARS Severe autism	29	5.78	28	.0000	-8.93	[-12.09; -5.77]	[-13.41; -4.45]	-6.06	-5.21

N: number of individuals; *df*: degree of freedom; d_{obs} *effect*: difference between P0 and P2[12]; *CI at 95% level*: Confidence Interval; *Pr (d > value) = 90%* or 95%: Bayesian probabilities at 90% and 95% for the *d* effect; *CrI at 95% level*: Bayesian Credible Interval.

As before, all CrIs are greater than CIs, and all *d* effects have a 50% probability of reoccurring for future data. For Global Developmental Ages (GDA), in 90% of cases, a 20-month difference may occur for children with mild to moderate autistic symptomatology, and a 14-month difference is expected for children with severe autistic symptomatology. For Verbal Developmental Ages (VDA), the gain is slightly higher for children with the mildest autistic symptomatology (22 months) and half as much for those with more severe autistic symptomatology (11 months). In terms of the degree of severity of autism (CARS), the most disturbed children could progress better than the others (about 6 points *vs* 2 points).

Overall, the CHIPPS program allows all children to improve their developmental profile in terms of overall intellectual or verbal abilities, and to decrease their autistic symptomatology, regardless of the degree of severity of autism or the intensity of intellectual disability at the beginning of the program implementation. Similarly, all children show a decrease in the intensity of their intellectual disability and the severity of their autistic symptomatology.

Using Bayesian inference, it was possible to predict progress for children with the same characteristics as the sample with different guarantees for 50%, 90%, 95%. And Credible Intervals provided a likely range that could be adjusted in different cases, taking into account the diversity of profiles.

Bayesian inference thus gives Psychopathology the possibility of finally being able to estimate in a valid and reliable way the magnitude of the effects of a program or treatment. In this study, it can therefore be argued that the CHIPPS program with LearnEnjoy applications will benefit children, regardless of their disability.

---

[12] Calculation direction: P2-P0.

     The second case study presented below does not include samples of a sufficiently large size to hope for generalization to a larger population. However, it is also possible to reliably quantify children's trends using other statistics adapted to "single cases".

5   Case Study 2: Psychological and Behavioral Evolution of ASD Children Following an ABA Program

The research aims to examine the psychological and behavioral evolution of children with autism or with Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS), who benefitted from an ABA intervention (Leaf, McEachin, & Taubman, 2010; Bernard-Paulais *et al.*, 2018) for 2 years. The Applied Behavior Analysis approach (ABA- Leaf *et al.*, 2010) was presented in Section 3.1 above. The research period includes three periods as previous study (P0 = beginning, P1 = after 1 year, P2 = after 2 years) corresponding to the different assessments of children with ASD receiving intensive care in school-based ABA. Psychological examinations and assessments of children's autonomy and coping behaviors are carried out by expert psychologists.

5.1   Method: Participants and Assessment Procedure

The sample consists of 10 children, named from *A* to *J*, with autism (n = 8) or other Pervasive Developmental Disorder – PDD - (n = 2), aged between 4 years 2 months and 11 years 1 month ($M = 6.4$; $SD = 2$) at the beginning of the study. Nine children presented with intellectual disability with different degrees of severity. All children benefitted from a full-time ABA intervention (Leaf *et al.*, 2010), implemented five days a week by two psychologists with a Board Certified Behavior Analyst-Doctorate (BCBA-D). The age of children's development is between 1 year 6 months and 4 years 5 months ($M = 2.6$; $SD = 1$). Children included in the study were previously diagnosed using DSM-IV-TR criteria (APA, 2000) and ICD-10 (OMS, 1999) and further diagnosed as presenting with an Autism Spectrum Disorder (ASD) based upon the DSM-5 (APA, 2003). The quantitative diagnostic evaluation of autism was carried out using CARS (Schopler *et al.*, 1988; Rogé, 1989). The diagnoses of ASD with or without ID were given by child psychiatrists and psychologists experienced in ASD and other neurodevelopmental disorders and blind to this study.

     As is often the case in psychopathology studies, it is extremely difficult to keep all participants on the scheduled time, and the different assessments could only be offered to 8 children, children *C* and *F* having left the study after one year.

     The SCEB (Social and Cognitive Evaluation Battery - Adrien, 2007; Adrien *et al.,* 2016; Thiébaut *et al.*, 2010) was administered to the 3 youngest children with a developmental age between 4 months and 24 months (children *D, E* and *I*).

     The EDEI-R (Differential Scales of Intellectual Efficiency-Revised - Perron-Borelli, 1997) was administered to 2 children (*B* and *H*) and the PEP-3 (Psycho-Educational Profile - Schopler *et al.*, 2010) to 4 children: *A, B, G, J*.

     We have selected for this study:
     a)   The results at the Differential Scales of Intellectual Efficiency-Revised (EDEI-R, Perron-Borelli, 1997) for both children *B* and *H*. These scales indicate the: Global Developmental Age (GDA), Verbal Developmental Age (VDA), Non-Verbal Developmental Age (NVDA), and Categorical Developmental Age (CDA) – recognizing categories/categorical processing.
     b)   The results at the PEP-3 for the children *A, B, G* and *J*.

     Considering the limited number of individuals assessed, the Single-case approach was applied in order to be able to finely and validly estimate children's behavioral and developmental progress.

5.2   Results at the Differential Scales of Intellectual Efficiency (EDEI-R): Children B and H

At the beginning of the study, the chronological age (CA) of *B* is 11 years 1 month and his Developmental Age (DA) is 4 years. For *H*, the CA is 5 years 6 months, and the DA is 4 years 5 months.

Children *B* and *H* show a favorable global evolution between periods P1 and P2 and then between P2 and P3, with an increase of 13 months for the GDA of child *B* between P1 and P3, and 14 months for child *H*. In the verbal domain of intellectual development (VDA), there is a heterogeneous evolution for each child with a very slight decrease between P2 and P3 for child *B* and a later progression (between P2 and P3) for child *H*. In the case of non-verbal development (NVDA), the evolution is more marked either between P1 and P2 for child *H* or between P2 and P3 for child *B*. Finally, a continuous progression of the categorical domain (CDA) is distinguished for both children with an evolution of 22 months between P1 and P3 for child *B*, and 34 months for child *H*, which is quite notable.

All domains combined, child *B* progresses on average by about 15 months in terms of DA between P1 and P3, and child *H* by 21 months (see Table 5 and Figure 6 hereafter).

Table 5. EDEI-R results: Individual effects, mean effects (M) and standard deviations (SD) for children B and H during the 3 test- periods (DA values in months)

	Child B			Child H		
	P2-P1	P3-P2	P3-P1	P2-P1	P3-P2	P3-P1
VDA	14	-9	5	2	11	13
NVDA	2	17	19	19	3	22
CDA	5	17	22	24	10	34
GDA	8	5	13	4	10	14
M	7.25	7.50	14.75	12.25	8.50	20.75
SD	5.12	10.71	6.50	10.90	3.20	8.41

In order to study more closely the evolution of children B and H, the Tau-U analysis presented above was implemented after visual inspection of the trends (Huskens *et al.*, 2012; Parker *et al.*, 2011).
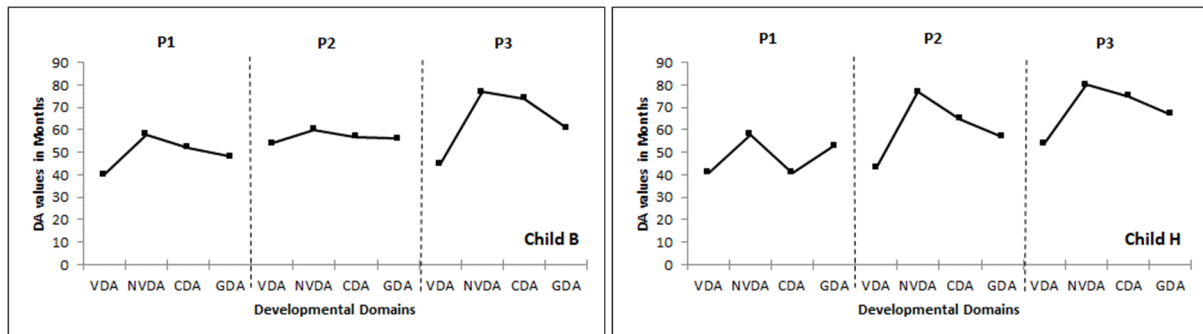

Figure. 6. Evolution of verbal (VDA), non-verbal (NVDA), categorical (CDA) and global (GDA) intellectual development obtained by children B and H at the 3 assessments' periods: P1, P2 and P3.

Visual inspection shows more overlap data for child B between periods P1 and P2 than for child H.

Tau_{novlap}, is an indicator of effect size (ES). Because it provides the proportion of all data pairs whose value could increase depending on the stages studied, it can be interpreted similarly to the Spearman Rhô or the Bravais-Pearson correlation coefficient. It is stronger here for Child H (Table 6) but there is a significant effect for the two children between the three periods concerned. Their progression is therefore strong, and we can conclude that the intervention has had a positive effect.

Table 6. Statistical summaries for children B and H for the 3 assessments periods P1, P2 and P3.

	Tau_{novlap}	p-value	Confidence Interval (CI) 90%
Child B	0.58	.020	0.1721<>0.9946
Child H	0.63	.012	0.1350<>1.000

5.3 Results at the Psycho-Educational Profile 3rd Edition (PEP-3): Children A, B, G and J

The items studied from PEP-3 concern three developmental assessments:
- Global evaluation: Communication (Com), Motor Skills (MSk), Global Development (GD),
- Communication: verbal/non-verbal cognition (Cog), Expressive Language (ExpLang), Receptive Language (RecLang) and
- Motor Skill: Fine Motor Skills (FinMSk), Global Motor Skills (GMSk) and Oculomotor Imitation (OcuImit).

There all are expressed in Developmental Ages (DA) values (in months). Figure 7 below, indicates the evolutions for the four children.
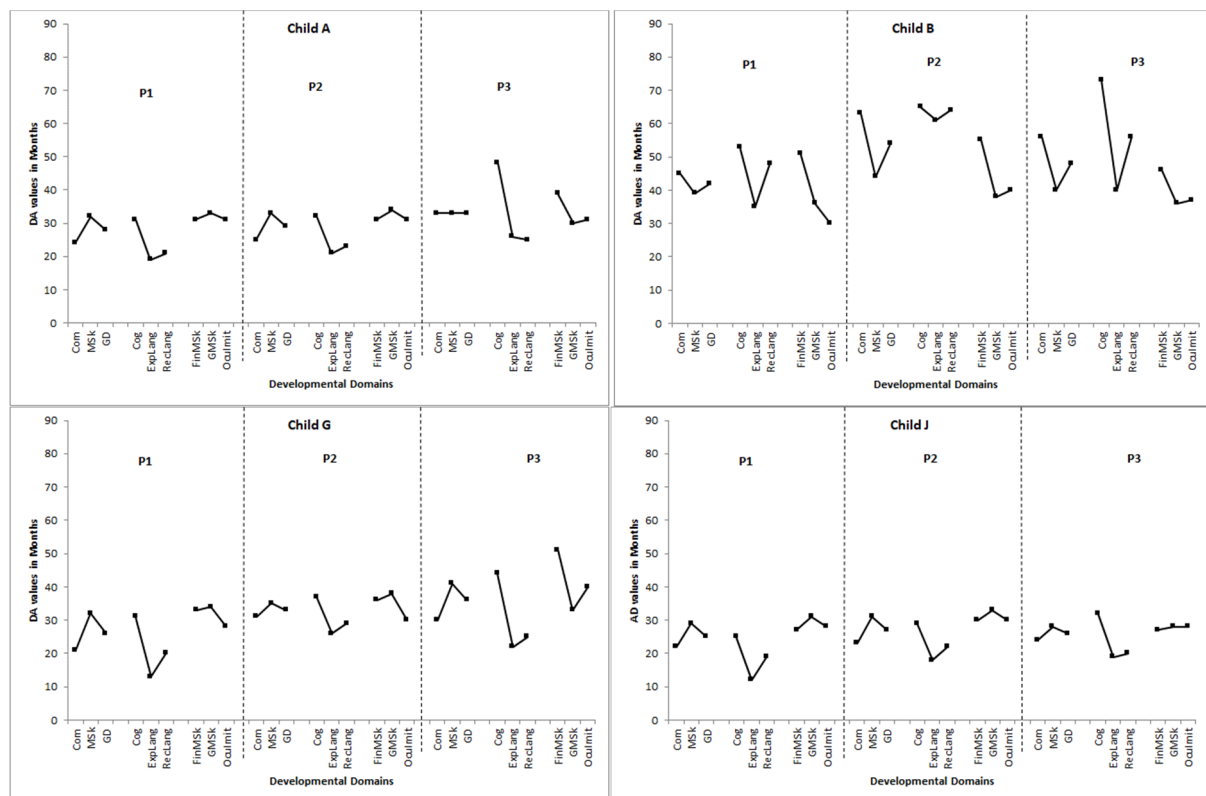

Figure. 7. Evolution of: Global evolution [Communication (Com), Motor Skill (MSk), Global Development (GD)], Communication [Verbal/non-verbal Cognition (Cog), Expressive Language (ExpLang), Receptive Language (RecLang)], Motor Skills [Fine Motor Skills (FinMSk), Global Motor Skills (GMSk) and Oculomotor Imitation (OcuImit)] obtained by children A, B, G and J at the 3 assessments periods: P1, P2 and P3.

Children's profiles are heterogeneous:
- Child A progresses on average by 5 months in terms of Global Development (Com, MSk and GD) and this effect is significant (Tau = 0.67; p =.02; CI 90%: 0.1831 <> 1.00), by about 9 months for cognition/language items (Cog, ExpLang, RecLang) and very little for items concerning motor skills (FinMSk, GMSk, OcuImit). These last two increases are not significant, given the significant overlap in data from P1 to P2.
- Child B improves his performance initially, regardless of the field, then his scores decline in P3 compared to those obtained in P2. Generally speaking, he obtains a gain of 6 months in Global Development and 11 months in Communication, but these results are not significant given the non-homogeneity of his results.
- Child G increases by about 9 months regardless of the families of items concerned after a slight decrease in P2 and the results are significant for Global Development items (Tau = 0.63; p =.03;

CI 90%: 0.2063 <> 1.00) and those concerning Motor Skills (Tau = 0.59; p =.04; CI 90%: 0.02 <> 1.00).

- Child J seems to regress to P2 and then P3 for Motor items, but he shows a 5-month increase in Cognition. The results are not significant because its profile is too heterogeneous.

Table 7 below summarizes the different effects of the period on these developmental ages according to each field.

Table 7. PEP-3 results: Individual effects, mean effects (*M*) and standard deviations (*SD*) for children A, B, G and J during the 3 test- periods (DA values in months)

		Child A			Child B			Child G			Child J		
		P2-P1	P3-P2	P3-P1	P2-P1	P3-P2	P3-P1	P2-P1	P3-P2	P3-P1	P2-P1	P3-P2	P3-P1
Global Eval.	Com	1	8	9	18	-7	11	10	-1	9	1	1	2
	MSk	1	0	1	5	-4	1	3	6	9	2	-3	-1
	GD	1	4	5	12	-6	6	7	3	10	2	-1	1
	M	1.00	4.00	5.00	11.67	-5.67	6.00	6.67	2.67	9.33	1.67	-1.00	0.67
	SD	0.00	3.27	3.27	6.51	1.53	5.00	3.51	3.51	0.58	0.58	2.00	1.53
Commu-nication	Cog	1	16	17	12	8	20	6	7	13	4	3	7
	ExpLang	2	5	7	26	-21	5	13	-4	9	6	1	7
	RecLang	2	2	4	16	-8	8	9	-4	5	3	-2	1
	M	1.67	7.67	9.33	18.00	-7.00	11.00	9.33	-0.33	9.00	4.33	0.67	5.00
	SD	0.58	7.37	6.81	7.21	14.53	7.94	3.51	6.35	4.00	1.53	2.52	3.46
Motor Skills	FinMSk	0	8	8	4	-9	-5	3	15	18	3	-3	0
	GMSk	1	-4	-3	2	-2	0	4	-5	-1	2	-5	-3
	OcuImit	0	0	0	10	-3	7	2	10	12	2	-2	0
	M	0.33	1.33	1.67	5.33	-4.67	0.67	3.00	6.67	9.67	2.33	-3.33	-1.00
	SD	0.58	6.11	5.69	4.16	3.79	6.03	1.00	10.41	9.71	0.58	1.53	1.73

To conclude, the results obtained from the intelligence and cognitive functioning tests show heterogeneous progress over time. All the children of this study have their own evolution defining several types of developmental trajectories: (a) a continuous developmental gain throughout the intervention, (b) a gain in the first period of the intervention then a slight decrease in the second period, (c) a developmental gain in the first period of the intervention then stabilization of cognitive performance in the second period.

These results obtained with the Single-Case method indicate the benefits of structured accompaniment based on the ABA program, over a period of 20 months, to children with autism and other pervasive developmental disorders. The benefits concern their autonomy, their behavior and the intensity of autistic symptomatology. Using this method, it is possible to better quantify individual performance and to be able to argue about the significance of progress.

6 Discussion and Conclusion

These two case studies are representative of the difficulties that psychologists involved in the psychopathology of autism regularly encounter: assessments are not always easy to administer with children's short-term attention problems and regulation disorders, there are defections during longitudinal studies, small samples of children with heterogeneous profiles, there is no possibility of creating a control group to judge the validity of support programs.

Bayesian and Single-Case methods provide valid and finer results than the simple description of data and allow to give consistency to interpretations.

These case studies analyzed with specific statistical methods offer two levels of understanding and interpretation of the results:

- A theoretical level with the objective of validating the models or theoretical foundations specific to intervention methods for people with autism.
- A practical and clinical level allowing to propose support integrative and naturalistic intervention programs scientifically tested in an ecological situation.

At the theoretical level:

The two case studies confirm the benefits of behavioral and developmental methods as well as the need to comply with the recommendations of the French High Authority of Health (HAS), if the wish is to allow this clinical population to progress in intellectual and behavioral domains.

At the clinical level:

- For the first case study, Bayesian method has shown an increase in intellectual abilities as well as a decrease in autistic behaviors for the 51 children with an Autism Spectrum Disorder benefiting from a global intervention program following in particular the HAS recommendations: the CHIPPS program. At the same time, new technologies as a complementary tool in remediation sessions provide undeniable benefits.

- For the second case study, the aim was not to study a sample, but individual trajectories. The statistical method was therefore different and more suited to the objectives. By analyzing the individual trajectories, it appeared that some children tended to regress during P2, and then progress was really visible in P3. This trend is not new for children with ASD. Indeed, in other studies conducted by Nézereau (2017) on resistance to change, there was also an increase in disorders after a few months of use of the support program, which then showed the beginning of stabilization and progress. Also, it would seem that people with ASD need time to adapt to new learning or to strengthen their cognitive abilities..

It is also important to highlight some aspects of these studies. Indeed, in recent years, several criticisms have been made about the multiplication of publications concerning interventions for people with ASD. For example, in her article, Chamak (2015) analyzes more than a hundred studies conducted with this population and summarizes her findings in these criticisms: (a) most research on the effectiveness of interventions proposes short-term protocols and is conducted on small samples and very young children. (b) Many studies are conducted with people with ASD, but frequently with intellectual levels that are within or close to normal. Thus, it is difficult to generalize the results to a larger population, given the heterogeneity of autistic disorders (Lombardo *et al.*, 2019).

With Bayesian or Single-Case methods, taking into account either large samples, heterogeneous in terms of symptoms, on a consistent longitudinal data collection and monitoring process, or small samples with significant heterogeneity, it can be shown that support programs are an undeniable help to children with ASD.

In conclusion, the results of these two case studies, analyzed using original, valid and rarely used statistical methods, validate the results of psychological monitoring, whatever the support method, and open up new perspectives for support, always with the idea of promoting the cognitive, social and autonomy development of people with ASD.

Conflict of interests: The authors have no conflict of interests to declare.

References

1. Adrien, J-L. (2007*). BECS, Batterie d'Évaluation Cognitive et Socio-émotionnelle.* Montreuil: Pearson- ECPA.
2. Adrien, J-L, Blanc, R., Thiébaut, E., Gattegno, M.P., Nader-Grosbois, N., Kaye, K., Pluvinage, C. Martineau, C., Clément, M., Machault, C., Hipeau-Bergeron, C-H., Bretière, M, Malandain, C., & Barthélémy, C. (2005). Etude préliminaire de la validation d'un nouvel instrument, la BECS (Batterie d'Evaluation Cognitive et Socio-émotionnelle), pour l'évaluation du développement psychologique d'enfants avec autisme. *Approche Neuropsychologique des Apprentissages chez l'Enfant (ANAE), 83-84*, 165-168.
3. Adrien, J.L., Martineau, J., Barthélémy, C., Bruneau, N., Garreau, B., & Sauvage, D. (1995). Disorders of regulation of cognitive activity in autistic children, *Journal of Autism and Developmental Disorders, 25, 3*, 247-261. doi: 10.1007/BF02179287.

4. Adrien J-L., Rossignol N., Martineau J., Roux S., Couturier G., and Barthélémy C. (2001). Regulation of cognitive activity and early communication development in young autistic, mentally retarded and young normal children. *Developmental Psychobiology, 39/2*, 124-136. doi: 10.1002/dev.1036

5. APA: American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders*: Fourth edition Text revision. Washington, DC: American Psychiatric Association.

6. APA: American Psychiatric Association. DSM-IV-TR (2003). *Manuel diagnostique et statistique des troubles mentaux: Texte révisé.* Paris: Masson.

7. APA: American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*: Fifth edition Text revision. Washington, DC: American Psychiatric Association.

8. Asperger, H. (1944). Die "Autistischen Psychopathen" im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten, 117,* 76-136. En ligne https://scholar.google.com/scholar_lookup?title=Die%20 Autistischen%20Psychopathen%20im%20Kindesalter&author=H..%20Asperger&journal=Archiv%20f%C3%BCr %20Psychiatrie%20udn%20Nervenkrankheiten&volume=117&pages=76-136&publication_year=1944.

9. Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London, 53,* 370–418. doi:10.1098/rstl.1763.0053.

10. Bernard, J.M. (1994). Introduction à l'analyse inductive. *Mathématiques, Informatique et Sciences Humaines, Numéro Spécial: Analyse des données planifiées, 126*, 71-80.

11. Bernard Paulais, M.A., Wolff, M., Nézereau, C., Bourgueil, O., Garnier, A., Collinet, M., Vandromme, L., & Adrien, J.-L. (2018). Etude de l'évolution cognitive, socio-émotionnelle et des comportements autistiques d'enfants avec autisme et autre trouble envahissant du développement bénéficiant d'une intervention structurée, la méthode Applied Behavior Analysis (ABA). *Annales Médico-Psychologiques, 176* (9), 847-856. doi: 10.1016/j.amp.2017.10.013.

12. Bernard Paulais, M-A, Mazetto, C., Thiébaut, E., Nassif, M-C, Costa Coehlo de Souza, M.T., Stefani, A.P., Blanc, R., Gattegno, M.P., Aïad, F., Sam, N., Belal, L. Fekih, L., Kaye, K., Contejean, Y., Wendland, J., Barthélémy, C., Bonnet-Brilhault, F. & Adrien, J-L (2019, June). Heterogeneities in cognitive and socio-emotional development in children with autism spectrum disorder and severe intellectual disability as a comorbidity. *Frontiers in Psychiatry, 10,* 1-14. doi: 10.3389/fpsyt.2019.00508

13. Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain, 128(2)*, 430–2441. doi: 10.1093/brain/awh561.

14. Blanc, R., Tourrette, C., Delétang, N., Roux, S., Barthélémy, C., and Adrien, J-L. (2000). Regulation of symbolic activity and development of communication in children with autism. *European Review of Applied Psychology, 50, 4*, 369-381.

15. Bölte, S., Golan, O., Goodwin, M. S., & Zwaigenbaum, L. (2010). What can innovative technologies do for autism spectrum disorders? *The National Autistic Society, 14*(3)*,* 155–159. doi: 10.1177/1362361310365028.

16. Bourgueil, O., Regnault, G., & Moutier, S. (2015). Création d'outils numériques pour personnes avec autisme: de la recherche à la pratique et vice versa. *Enfance, 1*, 111-126. doi: 10.4074/S001375451500107X.

17. Bourreau, Y., Roux, S., Gomot, M., Bonnet-Brilhault, F. & Barthélémy, C. (2009). Validation of the repetitive and restricted behaviour scale in autism spectrum disorders. *European child & adolescent psychiatry, 18 (*11), 675-682. doi: 1007/s00787-009-0028-5.

18. Cappe, E., Wolff, M., Bobet, R. & Adrien, J.-L. (2011). Quality of life: a key variable to consider in the evaluation of adjustment in parents of children with autism spectrum disorders and in the development of relevant support and assistance programs. *Quality of Life Research, 20(8),* 1279-1294

19. Casella, G., & Berger, L. (1987). Reconciling Bayesian and frequentist evidence in the one-sided testing problem (with discussion). *Journal of the American Statistical Association, 82*, 106-111.doi: 10.1007/BF02565110.

20. Chamak, B. (2015). Interventions en autisme: évaluations et questionnement. *Neuropsychiatrie de l'enfance et de l'adolescence,* 63, 297-301. doi: 10.1016/j.neurenf.2015.01.002.

21. Clark, K., & Green, G. (2004). Comparison of two procedures for teaching dictated-word/symbol relations to learners with autism. *Journal of Applied Behavior Analysis, 37,* 503–507. doi: 10.1901/jaba.2004.37-503.

22. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

23. Cohen, J. (1992). A power primer. *Quantitative methods in psychology, 112* (1), 155-159. doi: 10.1037/0033-2909.112.1.155.

24. Corroyer, D., & Wolff, M. (2003). *L'analyse statistique des données en psychologie: concepts et méthodes de base.* Paris: Armand Colin.

25. Degenne, C., Wolff, M., Fiard, D., & Adrien, J.-L. (2019). *ESAA: Evaluation Sensorielle de l'Adulte avec Autisme*. Paris: Hogrefe.

26. Delano, M. (2007). Improving written language performance of adolescents with Asperger syndrome. *Journal of Applied Behavior Analysis, 40*, 345–351. doi: 10.1901/jaba.2007.50-06.

27. Ferster, C.B., & DeMyer, M.K. (1961). The development of performances in autistic children in an automatically controlled environment. *Journal of Chronic Diseases, 13*(4), 312-345. doi: 10.1016/0021-9681(61)90059-5.

28. Fisher, R. A. (1935). The fiducial argument in statistical inference. *Annals of Eugenics. 5*(4), *391–398.* doi:10.1111/j.1469-1809.1935.tb02120.x. hdl:2440/15222.

29. Ganz, J.B., Hong, E.R., & Goodwyn, F.D. (2013). Effectiveness of the PECS Phase III app and choice between the app and traditional PECS among preschoolers with ASD. *Research in Autism Spectrum Disorders, 7,* 973-983. doi: 10.1016/j.rasd.2013.04.003.

30. Gattegno, M.P. (2004). L'accompagnement scolaire et professionnel des personnes atteintes d'autisme. School and professional coaching of people with autism. *Approche Neuropsychologique des Apprentissages chez l'Enfant (ANAE)*, *74-75*, 36-38.

31. Gattegno, M.P., Fernier, A., Granier-Deferre, C., & Adrien, J.L. (2005). Etude des effets de l'accompagnement d'enfants autistes à l'école et à domicile sur leur développement psychologique et social (Programme IDDEES). *Approche Neuropsychologique des Apprentissages chez l'Enfant, 83-84*, 196-202.

32. Gattegno, M.P., Wolff, M., & Adrien, J.L. (2012). Expérience française d'accompagnement scolaire en milieu ordinaire. In C. Philip, G. Magerotte, & J-L Adrien, *Scolariser des élèves avec autisme et TED* (pp. 75-91). Paris: Dunod.

33. Gattegno, M.P., Wolff, M., Ragonnet, P., Pascaud, R., de Fenoyl, C., & Adrien, J-L. (2017). Etude pilote de l'inclusion professionnelle en milieu ordinaire de personnes atteintes d'autisme. *Annales Médico-Psychologiques, 175*, 624-629. doi: 10.1016/j.amp.2016.06.009.

34. HAS: Haute Autorité de Santé. (2010). *Autisme et autres troubles envahissants du développement. États des connaissances hors mécanismes hors mécanisme physiopathologiques, psychopathologiques et recherche fondamentale.* Haute Autorité de Santé, Service Documentation – Information des publics.

35. En ligne https://www.has-sante.fr/portail/upload/docs/application/pdf/2010-03/autisme___etat_des_ connaissances_argumentaire.pdf.

36. HAS: Haute Autorité de santé. (2012). *Recommandation de bonne pratique: autisme et autres troubles envahissants du développement: interventions éducatives et thérapeutiques coordonnées chez l'enfant et l'adolescent.* Paris: Haute Autorité de santé. Service Documentation - Information des publics.

37. En ligne https://www.has-sante.fr/portail/upload/docs/application/pdf/2012-03/recommandations_autisme_ ted_enfant_adolescent_interventions.pdf.

38. Huskens, B., Verschuur, R., Gillesen, J., Didden, R., & Barakova, E. (2012). Promoting questions-asking in school-aged children with autism spectrum disorders: effectiveness of a robot intervention compared to a human-trainer intervention. *Journal of Developmental Neurorehabilitation, 16, 5*, 345-356. doi: 10.3109/17518423.2012.739212.

39. Jaynes, E.T. (2003). *Probability theory: the logic of science.* Cambridge, UK: Cambridge University Press.

40. Kamaruzaman, MF., Rani, NM., Nor, M., & Azahari, MH. (2016). Developing user interface design Application for children whith autism. *Procedia - Social and Behavioral Sciences*, *5*(217), 887-894.

41. Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2* (3) 217-250. En ligne https://pdfs.semanticscholar.org/f3d9/7cf1a54c7e0c8ee0be2e028f75ea0939e574.pdf?_ga=2.106232001.337058758. 1561151987-2050812248.1560503338.

42. Kendall, M.G., & Gibbons, J.D. (1999). *Rank correlation methods.* 5th edition. London: Arnold.

43. Leaf, R, & Mac Eachin, J. (2006). *Autisme et ABA, une pédagogie de progrès.* Paris: Pearson Education 2006

44. Leaf, R. B., Mac Eachin, J., Taubman, M., & Biesse, S. (2010). *L'approche comportementale de l'autisme. Bonnes et mauvaises pratiques, ce qu'il fallait en dire.* Paris: Pearson.

45. Lecoutre, B. (1998). From significance test to Fiducial Bayesian inference. In J.-M. Bernard, H. Rouanet, M.-C. Bert, B. Lecoutre, M.-P. Lecoutre, & B. Le Roux (Eds.), *New ways in statistical methodology – From significance tests to Bayesian inference* (pp. 123-157). Berne: Peter Lang.

46. Lecoutre, B., & Poitevineau, J. (1992). *PAC (Programme d'Analyse des Comparaisons).: guide d'utilisation et manuel de référence.* Montreuil: CISIA-CERESTA. En ligne http://www.modulad.fr/logiciels/lepac/33-ogiciel-pac/ErisLEPACPlus.html.

47. Le Roux, B., & Rouanet, H. (2004). Geometric data analysis: from correspondence analysis to structured data analysis. Bruxelles: Kluwer.

48. Lombardo, M.V., Lai, M.C., & Baron-Cohen S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Mol Psychiatr.* doi:10.1038/s41380-018-0321-0.

49. Martineau, J., Roux, S., Adrien, J.L., Garreau, B., Barthélémy, C., & Lelord, G. (1992). Electrophysiological evidence of different abilities to form cross-modal associations in children with autistic behavior. *Electroencephalographic Clinical Neurophysiology, 82,* 60-66.

50. Martineau, J., Adrien, J.L., Barthélémy, C., Garreau, B., & Lelord, G (1998). Association and regulation disorders in infantile autism. *Journal of Psychophysiology, 12,* 275-285.

51. Mesibov, G.B., Shea, V., & Schopler, E. (2005). *The TEACCH approach to autism spectrum disorders.* New York: Springer Publishers.

52. Munoz, R., Barcelos, T., Noel, R., & Kreisel, S. (2012). Development of Software that Supports the Improvement of the Empathy in Children with Autism Spectrum Disorder. *In Proceedings of the the Chilean Computer Science Society (SCCC), IEEE* (pp. 223-228). Valparaiso, Chile. doi: 10.1109/SCCC.2012.33.

53. Murphy, C.A. (2006). *The comparative effects of simple and complex instructional language on the acquisition and generalization of receptive language tasks by children with autism.* PHD-thesis. Ohio University. En ligne https://etd.ohiolink.edu/.

54. Nézereau, C. (2017). *Étude des trajectoires développementales d'enfants et d'adultes avec Trouble du Spectre de l'Autisme bénéficiant du programme IDDEES et incluant de nouvelles technologies d'apprentissage scolaire et de communication (LearnEnjoy).* Thèse de Doctorat. Université Paris Descartes.

55. Nézereau, C., Wolff, M., Gattegno, M.P., Adrien, J-L. (2018). Évaluation ergonomique d'applications mobiles (LearnEnjoy) dédiées à l'apprentissage scolaire et la communication pour les personnes atteintes du Trouble du Spectre de l'Autisme. In *Proceedings of the Ergo'IA 2018 conference.* New-York: ACM. En ligne https://www.researchgate.net/profile/Marion_Wolff/publication/329177743_Evaluation_ergonomique_d%27applications_mobiles_LearnEnjoy_dediees_a_l%27apprentissage_scolaire_et_la_communication_pour_les_personnes_atteintes_du_Trouble_du_Spectre_de_l%27Autisme/links/5bfab2e5299bf1a0203232a3/Evaluation-ergonomique-dapplications-mobiles-LearnEnjoy-dediees-a-lapprentissage-scolaire-et-la-communication-pour-les-personnes-atteintes-du-Trouble-du-Spectre-de-lAutisme.pdf.

56. Nézereau, C., Wolff, M., Gattegno, M.P., & Adrien, J-L. (2020). Evaluation ergonomique d'applications numériques dédiées aux personnes atteintes du Trouble du Spectre de l'Autisme (TSA) pour l'apprentissage scolaire et la communication. In M. Wolff & R. Mollard (Eds.), *Pratiques de l'ergonomie : de la méthode aux applications* (pp. 245-271). Toulouse, F: Octarès.

57. Nézereau, C., Wolff, M., Gattegno, M-P., Bourgueil, O., Regnault, G., & Adrien, J-L**.** (2016). Evolution de la régulation et de la résistance au changement d'enfants et d'adultes avec Trouble du Spectre de l'Autisme (TSA): contribution des applications numériques «LearnEnjoy» dans le cadre d'un programme d'intervention développementale, le programme IDDEES. In *Proceedings of the Ergo'IA 2016 conférence* (6-8). New-York: ACM. doi: 10.1145/3050385.3050398.

58. Organisation Mondiale de la Santé (1999). CIM-10 Classification statistique internationale des maladies et des problèmes de santé connexes. 10ᵉ révision. Genève: OMS. En ligne https://www.atih.sante.fr/sites/default/files/public/content/2665/cim10_2015_final_0.pdf.

59. Ozonoff, S. (1997). Components of executive function in autism and other disorders. In James Russell (Ed). Autism as an executive disorder. (pp. 179-211). New York, NY, US: Oxford University Press.

60. Parker, R.I., Vannest, K.J., & Davis, J.L. (2011). Effect size in single-case research: a review of nine nonoverlap techniques. *Behavior Modification, 35,* 303-322. doi: 10.1177/014544551139914.

61. Parker, R.I., Vannest, K.J., Davis, J.L., & Sauber, S.B. (2011). Combining nonoverlap and tred for single-case research: Tau-U. *Behavior Therapy, 42,* 284-299. doi: 10.1016/j.beth.2010.08.006.

62. Perron-Borelli, M. (1997). *Echelles différentielles d'efficiences intellectuelles, forme révisée (EDEI-R).* Paris: Pearson-Editions du Centre de psychologie appliquée.

63. Poitevineau, J. (2004). L'usage des tests statistiques par les chercheurs en psychologie: aspects normatifs, descriptifs et prescriptifs. *Math. & Sc. Hum., 167, 3,* 5-25. doi: 10.4000/msh.2900.

64. Reichow, B., Volkmar, F.R., & Cicchetti, D.V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders, 38, 7,* 1311-1319. doi: 10.1007/s10803-007-0517-7.

65. Robert, C. (2013). *Des spécificités de l'approche bayésienne et de ses justifications en statistique inférentielle.* HAL Id: hal-00870124. En ligne https://hal.archives-ouvertes.fr/hal-00870124.

66. Rogé, B. (1989). *Echelle d'évaluation de l'autisme infantile (CARS).* Issy-les-Moulineaux: Editions d'Applications psychotechniques.

67. Rogers S., & Dawson, G. (2013). L'intervention précoce en autisme: le modèle de Denver pour jeunes enfants. Évaluation et prise en charge. Paris, Dunod.

68. Rouanet, H. (1996). Bayesian methods for assessing importance of effects. *Psychological Bulletin, 119*, p. 149-158. doi: 10.1037/0033-2909.119.1.149.

69. Rouanet, H., Bernard, J.-M., Bert, M.-C., Lecoutre, B., Lecoutre, M.-P., & Le Roux, B. (1998). *New ways in statistical methodology.* NY: Peter Lang.

70. Sani-Bozkurt, S., Vuran, S., & Akbulut, Y. (2017). Design and Use of Interactive Social Stories for Children with Autism Spectrum Disorder (ASD). *Contemporary Educational Technology*, *8*(1), 1-25. En ligne https://files.eric.ed.gov/fulltext/EJ1126817.pdf.

71. Schreibman, L., Dawson, G., Stahmer, A.C., Landa, R., Rogers, S.J., McGee, G.G., Kasari, C., Ingersoll, B. Kaiser, A.P., Bruinsma, Y., McNerney, E. Wetherby, A., Halladay, A. (2015). Naturalistic Developmental Behavioral Interventions: Empirically validated treatments for Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 45:2411–2428 DOI 10.1007/s10803-015-2407-8.

72. Schopler, E., Reichler, RJ. & Lansing, M. (2002). Stratégies éducatives de l'autisme. Paris: Masson.

73. Schopler, E., Reichler, R.J. & Renner, B.R. (1988). *The Childhood Autism Rating Scale (CARS).* Los Angeles: Western Psychological Services.

74. Scruggs, T.E., & Mastropieri, M.A. (1994). The utility of PND statistic: a reply to Allison and Gorman. *Behavior Research and Therapy, 32,* 879-883. doi: 10.1016/0005-7967(94)90169-4.

75. Scruggs, T.E., Mastropieri, M.A., & Casto, G. (1987). The quantitative synthesis of single subject research: methodology and validation. *Remedial and Special Education, 8, 2*, 24-33. doi: 10.1177/074193258700800206.

76. Taubman, M., Brierley, S., Wishner, J., Baker, D., McEachin, J., & Leaf, R.B. (2001). The effectiveness of a group discrete trial instructional approach for preschoolers with developmental disabilities. *Research in Developmental Disabilities, 22*(3), 205-219.

77. Vannest, K.J., Parker, R.I, & Gronen, O. (2011). *Single case research: web based calculators for SCR analysis (version 1.0).* [Web-based application]. College station, TX: Texas A&M University Available from singlecaseresearch.org (retrieved).

78. Verhagen, J., & Wagenmakers, E.J. (2014). Bayesian tests to quantify the result of a replication attempt. Journal of Experimental Psychology, *143* (4), 1457–1475. doi: 10.1037/a0036731.

79. Wechsler, D. (2014). *Wechsler intelligence scale for children-fifth edition.* Bloomington, MN: Pearson.

80. Wolff, M. (2003). Apports de l'analyse géométrique des données pour l'analyse de l'activité. In J.-C. Sperandio & M. Wolff (Eds.), *Formalismes de modélisation pour l'analyse du travail et l'ergonomie* (pp. 195-227). Paris: PUF.

81. Wolff, M., Gattegno, M.P., & Adrien, J.L**.** (2008). Accompagnement et intégration de personnes avec autisme: étude du rôle du psychologue superviseur par les techniques conjointes d'entretien et d'analyse géométrique des données. In D. Galarreta, P. Girard, J.-C. Tucoulou, & M. Wolff (Eds.), *L'humain au cœur des systèmes et de leur développement: Ergo'IA, 2008* (pp. 71-78). Biarritz, France: Estia Innovation.

82. Wolff, M., Gattegno, M.P., Adrien J.L, Gabeau, C., & Isnard, P. (2014). Contribution of tablets to the support of children and adolescents with autistic disorders. *European Journal of Automation, (EJA), 4-5-6*, 261-282. doi:10.3166/jesa.48.261-282.