COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

AUTISM SUBTYPES IN CONTEXT OF CURRENT THEORIES

Diploma Thesis

Bratislava, 2017 Bc. Nikola Ondríková

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

AUTISM SUBTYPES IN CONTEXT OF CURRENT THEORIES

Diploma Thesis

Study programme:	Cognitive Science
Field of study:	2503 Cognitive Science
Supervisor:	Prof. MUDr. Daniela Ostatníková, PhD.
Consultant:	Mgr. Hana Celušáková

Bratislava, 2017 Bc. Nikola Ondríková

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname: Bc. Nikola Ondríková
Study programme: Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study: Cognitive Science
Type of Thesis: Diploma Thesis
Language of Thesis: English
Secondary language: Slovak

Title: Autism Subtypes in Context of Current Theories

Aim: Illustrate autism spectrum disorders in the context of current theories. Create subtypes of autism with different clustering methods, and/or their combinations. Discuss results in the context of current theories of autism.

Literature:
Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. The computer journal, 41(8), 578-588.
Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian theories of autism spectrum disorder Help improve Clinical practice?. Frontiers in psychiatry, 7.
Shen, J. J., Lee, P. H., Holden, J. J., & Shatkay, H. (2007). Using cluster ensemble and validation to identify subtypes of pervasive developmental disorders. In AMIA Annual Symposium Proceedings (Vol. 2007, p. 666). American Medical Informatics Association.

Annotation: Autism is a neurodevelopmental disorder with effects like stereotypical behaviour and difficulties in communication. In practice, clinicians use classification from ICD-10 which is lately perceived as inaccurate. Assigning the right subtype can be crucial for future prediction concerning the development of disorder-specific symptoms.

Supervisor: prof. MUDr. Daniela Ostatníková, PhD.
Consultant: Mgr. Hana Celušáková
Department: FMFI.KAI - Department of Applied Informatics
Head of department: prof. Ing. Igor Farkaš, Dr.

Assigned: 20.06.2016

Approved: 22.11.2016 prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

.....
Student

.....
Supervisor

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Nikola Ondríková
Študijný program: kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: kognitívna veda
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Autism Subtypes in Context of Current Theories
Podtypy autizmu v kontexte súčasných teorií

Cieľ: Popísať autizmus v kontexte aktualnych teorii. Vytvoriť podtypy autizmu pomocou rôznych zhlukových analýz a ich kombinácií. Zhodnotiť výsledky v zmysle súčasných teorii a diskutovať o možných aplikáciach.

Literatúra:
Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. The computer journal, 41(8), 578-588.
Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian theories of autism spectrum disorder Help improve Clinical practice?. Frontiers in psychiatry, 7.
Shen, J. J., Lee, P. H., Holden, J. J., & Shatkay, H. (2007). Using cluster ensemble and validation to identify subtypes of pervasive developmental disorders. In AMIA Annual Symposium Proceedings (Vol. 2007, p. 666). American Medical Informatics Association.

Anotácia: Autizmus je vývinová porucha nervového systému s následkami ako napríklad stereotypné správanie alebo problémy v komunikácií. V praxi sa používa klasifikácia založená na MKCH-10, ktorá je v poslednej dobe vnímaná ako nepresná. Priradenie správneho podtypu môže byť zásadné vzhľadom na predikciu budúcich symptómov.

Vedúci: prof. MUDr. Daniela Ostatníková, PhD.
Konzultant: Mgr. Hana Celušáková
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, Dr.

Dátum zadania: 20.06.2016

Dátum schválenia: 22.11.2016

prof. Ing. Igor Farkaš, Dr.
garant študijného programu

.....
študent

.....
vedúci práce

Declaration

I hereby declare that the work presented in this thesis is original and result of my own investigations. Formulations and ideas taken from other sources are cited as such.

Bratislava, 2017

Acknowledgement

My thesis benefited from the valuable consultations with Prof. MUDr. Daniela Ostatníková PhD. and Mgr. Hana Celušáková from Center for Autism Research at Comenius University in Bratislava.

ABSTRACT

Autism is a neurodevelopmental disorder usually defined with a triad of manifestations – 1) impaired social interaction and 2) communication, and 3) repetitive behavior. The question what causes the symptomatology remains unanswered. The thesis aims to describe autism in a context of traditional theories of autism. Bayesian Brain Hypothesis is currently predominant theory of autism in cognitive science. The theory offers a roof for most of the other theories. Bayesian Brain Hypothesis understands the brain as a probabilistic machine which uses its general model to generate predictions. Current sensory input is tested against predictions or priors, and general internal model updates its beliefs about causes of the input. Predictions of the general model are inaccurate in autism. Bayesian perspective is promising account for autism subtyping due to possibility of modelling subtypes using various parameters. Future computational models could help to better understand the nature of the differences within autistic group and explain inconsistencies in biological research. From previous findings, it seems like there are two main subtypes - "less severe" and "more severe". In this study, we derive clusters from behavioral data employing cluster validation and hierarchical cluster analysis with model-based clustering. Our dataset consists of children's behavior description from their parent's standardized Interview (ADI-R) and clinician's observation (ADOS). Sample involves 217 autistic children, 13,36% girls 7,2 years old in average (SD =4,44) and 86.64% boys 7,5 (SD = 4,64) years old in average. Our results are in line with earlier findings and suggest the existence of two main clusters based on onset of developmental anomalies and gesture communication. We present a possibility of association between severity of autism, ability to form representations and motivation to social stimuli.

KEYWORDS: autism, Bayesian Brain hypothesis, cluster analysis

ABSTRAKT

Autizmus je porucha neuronálneho vývinu, zvyčajne definovaná trojicou prejavov - 1) zhoršená sociálna interakcia a 2) komunikácia a 3) repetitívne správanie. Otázka, čo spôsobuje symptomatológiu, zostáva nezodpovedaná. Cieľom práce je popísať autizmus v kontexte tradičných teórií autizmu. Hypotéza Bayesovského mozgu v súčasnosti prevláda teóriou autizmu v kognitívnej vede. Teória ponúka strechu pre väčšinu ostatných teórií. Hypotéza Bayesovského mozgu chápe mozog ako pravdepodobnostný stroj, ktorý používa svoj všeobecný model na generovanie predpovedí. Aktuálny senzorický vstup je testovaný voči predpovediam a všeobecný interný model aktualizuje svoje presvedčenia o príčinách vstupu. Predpovede všeobecného modelu sú v prípade ľudí s poruchou autistického spektra nepresné. Bayesovská perspektíva je sľubná aj v rámci vytvárania podtypov autizmu. Najmä kvôli možnosti modelovania podtypov pomocou rôznych parametrov. Vypočtový model by mohlo pomôcť lepšie pochopiť podstatu autizmu a vysvetliť nezrovnalosti v biologickom výskume. Z predchádzajúcich zistení sa zdá, že existujú dva hlavné podtypy - "menej závažný" a "zavažný". V práci sme vytvorili podtypy autistov na základe behaviorálnych údajov, s použitím metód validácie klastrov. Ďalej sme použili hierarchickú zhlukovú analýzu a zhlukovú analýzu založenú na modelovaní. Behaviorálne údaje pozostávajú z opisu správania detí z rodového štandardizovaného rozhovoru (ADI-R) a pozorovania klinického psychológa (ADOS). Výskumná vzorka zahŕňa 217 autistických detí, 13,36% dievčat v priemernom veku 7,2 (SD = 4,44) a 86,64% chlapcov v priemernom veku 7,5 (SD = 4,64). Naše výsledky sú v súlade s predchádzajúcimi zisteniami a naznačujú existenciu dvoch hlavných klastrov založených na vzniku vývojových anomálií a komunikácii gestami. Predstavujeme možnosť prepojenia medzi závažnosťou autizmu, schopnosťou vytvárať reprezentácie a motiváciou k sociálnym stimulom.

KĽÚČOVÉ SLOVÁ: autizmus, Hypotéza Bayesovského mozgu, klastrová analýza

CONTENT

Introduction

The thesis aims to investigate nature of autism and its subtypes. Autism is considered as a neurodevelopmental disorder with spectral character. The spectrum is however very broad, and research samples consisting of very different cases can result in conflicting findings of biological studies. Autism is associated with cognitive theories focusing on "local" way of thinking like Extreme Male Brain theory (Baron-Cohen, 2002) or Weak Central Coherence hypothesis (Happé, 1999). Authors speculate about autism as about cognitive style (i.e. a way of thinking) because of its spectral nature. The theories describe an autistic cognitive style as very detail-focused and systematic. These concepts, however, have not provided an elaborate explanation of impairments in autism and cannot explicate possible subtypes. Quite recently, considerable attention has been paid to Bayesian brain hypothesis which should be the one box for all previously mentioned views. The reason that Bayesian perspective is likely to account for autism subtyping is a possibility to computationally model subtypes using various parameters. Additionally, Bayesian Brain hypothesis approach symptomatology in an interactive manner and therefore explain nature of autism better (Haker, Schneebeli & Stephan, 2016).

Former research in autism subtyping suggested at least four clusters (Beglinger, & Smith, 2001). Current research has conversely shown (Hu & Sternberg, 2009; Veatch et al., 2014) that, there are two main subtypes - "less severe" and "more severe". The two subtypes define extremes of continuum in terms of severity. Clusters between the extremes vary from research to research. A key limitation of autism subtyping studies is that analyses are based on various individuals, usually diagnosed just with one diagnostic method. This could cause inclusion of syndromic autistic children in which autism is a secondary consequence of genetic syndrome and misdiagnosed children with language deficits, ADHD, etc..

The Bayesian Brain hypothesis is predominant in current literature concerning autism; therefore, we focus on this theory in greater extent in comparison with other theories. Concerning structure, we start with a brief description of traditional theories, then we concentrate on the Bayesian Brain in Autism, and we summarize research in behavioral and sensory subtypes. In the next part, we characterize methods and results follows. In the discussion, we discuss our result in the context of theories of autism and in the context of autism research in general. Thesis combines literature from psychology, neuroscience,

genetics and applies supervised learning to analyze received data. We approach the topic in an interdisciplinary manner.

1 Autism

Autism is a neurodevelopmental disorder usually defined with the triad of manifestations – impaired social interaction and communication, and repetitive behavior. Diagnosis is usually based on the triad which is understood as follows (Thorová, 2006): *1) Disruption of mutual social interaction* takes the form of inadequate evaluation of social-emotional situations, as reflected inadequate response to other people's emotions or lack of adapting the behavior of social context, misuse of social signals, weak integration of social, emotional and communicative actions or lack of social-emotional reciprocity; *2) Impaired quality of communication* - It takes the form of a limited social use of language (regardless of the level of language skills), failure of the imagination and socially mimetic game, the lack of synchronization and the lack of reciprocity in the communications conversation, reduced adaptability of language expression and a relative lack of creativity and fantasy in thought. Lack of emotional response to the friendly approach of other people, whether it is verbal or non-verbal, this impaired use causes changes in cadence or emphasis that reflect the modulation of communication and insufficient gestures to underline the importance of spoken communication; *3) Restricted, repetitive and stereotyped behaviors, interests and activities* - The tendency to rigidity and routine behavior in a wide range of aspects of everyday life. There may be a particular attachment to an unusual object or its parts. Children can take on carrying out routines and rituals of a non-functional character. It is often movement or interest in non-functional elements of objects (e.g. The smell or surface). A child may be resistant to changes in the ordinary course of everyday activities and the details of the personal environment (e.g. The movement of decorations or furniture in the family house).

This triad of symptoms is still used in ICD-10 (WHO, 1993), DSM-5 (APA, 2013) present different view though. Autism used to be part of the bigger category – Pervasive Developmental Disorder, often abbreviated as PDD. Autism with Asperger syndrome and Pervasive Developmental Disorder – not otherwise specified form a new diagnosis – Autism Spectrum Disorder (ASD). To sum up, instead of three diagnosis there is only one. Removing categories, however, does not mean there are no subtypes at all. With the diagnosis of ASD goes specification depending on language ability, intelligence and various comorbidity typically occurring with ASD. The reason for such change can be explained by a heterogeneous view of mental disorders in general. “The historical aspiration of achieving diagnostic homogeneity by progressive subtyping within disorder categories no longer is

sensible; like most common human ills, mental disorders are heterogeneous at many levels, ranging from genetic risk factors to symptoms" (APA, 2013). They also changed the triadic view of symptoms to the duo of indications: 1) Social Communication and Interaction and 2) Fixed Interests and Repetitive Behaviors. Lai, Lombardo, Chakrabarti & Baron-Cohen (2014) appreciate the changes because they will probably improve the efficiency of clinical and educational centers for ASD individuals. Nonetheless, as authors pointed out, diagnosis of adults and milder forms can be more difficult because of compensatory strategies and research can get more complicated as well. Several biological roots of autism were identified (Chaste & Leboyer, 2012) and therefore there should be subtypes.

1.1 Theories of Autism

In this part, we will describe current theories (i.e. still present in current research) of autism. From different areas of the research come various attempts to explain nature of autism but we still miss the whole picture. We start chronologically from cognitive theories, and we end with neurobiological theories or hypothesis. To best of our knowledge, we pick up the most common topics in autism research even.

One of the oldest theories in autism research is "Weak Central Coherence" (WCC) theory (Frith & Happé, 1994; Happé, 1999). WCC theory suggests that cognitive impairment in ASD is caused by the inability to see the whole picture. Regarding the theory, autistic individuals are focused on are focused on detailed information about objects omitting a context rather than perceive complex environment settings. Happé (1999) proposes to understand WCC as a local-global cognitive style. Behind this proposal is an assumption that people with autism have the specific way of thinking emphasizing local relationships and associations present in the healthy population. To further investigate this idea, Happé, Frith & Briskman (2001) conducted a study comparing WCC cognitive style in fathers of dyslectic, autistic and typically developing boys. Fathers of autistic children were biased in the same manner as their children favoring detail information processing over processing of the whole concept. By contrast to their children, this way of thinking is advantageous. Despite the popularity of this approach, its claims, and some supporting empirical evidence (Vanegas & Davidson, 2015) the theory is continually unsuccessful in providing sufficient general explanation of information processing and attentional focus in ASD (Pellicano, 2012). It also does not explain why autistic people would be focused on details (Baron-Cohen, 2002).

Following theory has wider explanatory power. Baron-Cohen (2002) uses the concept of dimensional cognitive style – empathizing / systemizing, to explain different types of "brain". One end of the continuum is empathizing characterized by understanding social world and relationships, and it is typical for "Female Brain". This ability is often called Theory of Mind. On the other end of the continuum is systemizing portrayed as "Male Brain". A person with Male Brain has strong analytical and technical skills, and he or she has high sensitivity to details. Concerning autism, he proposes "Extreme Male Brain Theory" (EMBT). The EMBT suggest an extremely low level of empathizing skills and strong systemizing abilities. In other words, a person with such a brain type would have a serious problem with mind reading, the proper reaction in a new social environment and social interaction in general. Instead, this person would excel in systemic domains like organizing items, mechanics or another rule-based subject. Autistic individuals are focused on details because they want to understand the closed systems in general. However, if there is the extreme male brain, there probably should be extreme female brain as well. Larson et al. (2015) conducted a study examining autistic individuals with psychotic comorbidity and found empathizing bias. This result indicates eligibility of the EMBT. From the biological point of view, it is not clear. The theory is biologically connected with fetal testosterone exposure (Teatero & Netley, 2013). A meta-analysis (Teatero & Netley, 2013) reviewing studies investigating the main biomarker of fetal testosterone exposure supported the EMBT partially but identified specific issues. Nevertheless, it is problematic to say whether the problem is due to theory or quality of the biomarker. We also have a poor understanding of the link between prenatal hormones exposure and specific behavioral traits (Whitehouse, 2016). The theory is neuroscientifically connected to the amygdala, mostly because of mind reading issues and lack of social interest (Baron-Cohen et al., 2000). The current review of amygdala theory of autism (Zalla & Sperdutti, 2013) has proposed that primary deficit may be a diminished ability to orient attention towards significant stimuli caused by an insult to the interconnected fronto-amygdala circuit. Social motivation theory can explain these findings more consistently.

According to social motivation theory (Chevalier, Kohls, Troiani, Brodkin & Schultz, 2012), the social world directs our attention. Attention prioritizes social signals because interactions are highly rewarding. Social motivation is based on biological mechanisms. Thus, can be understood as an evolutionary advancement to collaborative environment typical for humans. Enhancing attention to social information and promoting the desire to maintain social bonds upgrades relationships and encourages coordination and

collaboration. In contrast, autistic individuals appear to have an overall decrease in value attributed to social information. Known deficits in face processing in persons with an autism correlate with a lack of a fusiform face area (Grelotti, Gauthier & Schultz, 2002). Because individuals with ASD have probably decreased social interest, they may fail to form cortical face specialization. Possibly disrupted biological roots are the orbitofrontal-striatal-amygdala circuitry as well as malfunction of certain neurotransmitters. Theory sees autism as an extreme case of early-onset diminished social motivation (Chevalier, Kohls, Troiani, Brodkin & Schultz, 2012; Grelotti, Gauthier & Schultz, 2002).

Research in structural changes and head circumference is not so clear as it may seem from amygdala theory. Results indicate either decreased (e.g. Freitag et al., 2009) or increased overall brain size (Fidler, Bailey & Smalley, 2000). To tangle the problem, even more, there is evidence for normal head circumference stated by Aylward et al. (1999). Meta-analysis of results (Redcay & Courchesne, 2005) has shown that the deviation from normal is mainly prevalent between two and five years of age. The method of measuring head circumference has also had a significant impact on results. Regarding structural changes, neuroimaging research indicates abnormalities in the volume of the amygdala (Baron-Cohen, Bullmore, Wheelwhright, Ashwin & Williams, 2000), gray matter volume (Palmen et al., 2005), cortical thickness (Zielinski et al., 2014) and lot more. Empirical evidence also suggests these changes differ with aging (Courchesne, Campbell & Solso, 2011; Zielinski et al., 2014). It is worth to mention that such neuroimaging studies are usually done on small samples of individuals with high-functional autism or Asperger syndrome.

Speaking of autistic brain, another essential topic in the neuroscience of autism is brain connectivity. As usually, research offers opposing findings. There are studies supporting under-connectivity (e.g. Martino et al., 2014) and, on the other hand, studies encouraging over-connectivity of the autistic brain (e.g. Müller et al. 2011). In 2015, Nature published resting-state brain study authored by Hahamy, Behrmann & Malach, wherein this problem is gripped. The inter-hemispheric connectivity for different areas of the neurotypical brain seems to be standardized. In ASD individuals, reduced connectivity was found in the frontal and temporal cortices responsible for executive decision making. On the other hand, connectivity is increased in the occipital and sensorimotor cortices, where sensory and motor processing mostly occurs. Moreover, “scan after scan” showed that neurotypical brains exhibited uniform patterns in relative connectivity. It was harder for the autistic group. While all of them had some areas of over- or under-connectivity, the regions

in which these occurred, individually varied. When scans were judged as a group, there were many patterns, but when individual scans were compared with ASD symptomatology, significant correlations were observed. Authors propose that individualized changes in functional connectivity organization are a core characteristic of high-functioning ASD and that this can be the solution for contradictory findings.

2 Autism from Bayesian Perspective

Bayesian Brain hypothesis (Knill, & Pouget, 2004) is a theory which understands the brain as a probabilistic machine which uses its general model to generate predictions. Current sensory input is tested against predictions, and general internal model updates its beliefs about causes of the input. How the prediction update is described with free energy principle or predictive coding (Friston, 2010) what are different names for ideas originate in Kant and were later redefined by Hermann von Helmholtz (Friston, 2016; Swanson, 2016).

2.1 Cognition from Bayesian Perspective

To better comprehend how Bayes' Theorem can help us understand cognition in general, we follow explanation outline from Haker, Schneebeli, & Stephan (2016). These authors explicate Bayesian Brain thought four concepts: (1) Learning, (2) Uncertainty, (3) Cognitive Hierarchies, and (4) Homeostasis.

Psychology defines learning in many ways. The authors defining learning from a psychological point of view, usually try to capture all related factors. However, in computational sciences, it is better to have more concise characterization. In this thesis, to learn something means to get new information and incorporate or interiorize this information to gain knowledge. In Bayesian Brain, the process of learning is described by predictive coding and free energy principle. Both concepts describe the same mechanism (Friston, Kilner, & Harrison, 2006). To simply summarize the conclusion of Friston, Kilner & Harrison (2006), the goal of a self-organizing system in balance with its environment is to minimize prediction error (according to predictive coding) or free energy (according to free-energy principle). This free energy is not a consequence of thermodynamic processes but biologically speaking, it emerges from dynamics and selection in the population of neurons or animals. The lower the prediction error or the free energy, the better.

This goal can be fulfilled by (1) action – "moving away" or (2) learning – "explaining away" (Haker, Schneebeli, & Stephan, 2016). Satisfying prediction by action means to act according to prediction; then the prediction error is zero. In the situation of incorrect prediction, there is a prediction error which is used to improve the internal model. During this update process, the prediction error is explained away.

In “A Bayesian foundation for individual learning under uncertainty” Mathys, Daunizeau, Friston, & Stephan (2011) rewrite Bayes’s theorem like this:

$$\Delta\text{belief} \propto \frac{\text{precision}_{\text{input}}}{\text{precision}_{\text{prior belief}}} \times \text{prediction error}$$

It means that “any change in belief is proportional to prediction error, but weighted by the ratio of the precision of the sensory input and the precision of the prior belief” (Haker, Schneebeli, & Stephan, 2016, p. 5). We can look at the precision ratio as to dynamic learning rate. Particularly, every time the confidence in the sensory input is higher than the confidence in the current belief, the learning rate is high. In other words, when the uncertainty of the internal prediction is larger than the uncertainty about sensory input we change something in our general internal model (Haker, Schneebeli, & Stephan, 2016).

The human brain and process of Bayesian inference have a hierarchical structure. Therefore, we can understand the brain with the analogy of Bayesian inference. Besides anatomical hierarchy of the brain, there is also functional hierarchy which is related to long-term networks. According to hierarchical clustering of human functional brain networks, it has two main branches (Gleiser & Spoormaker, 2010): the higher parts consist of subcortical, limbic and paralimbic regions and the lower parts from the neocortical association and sensory cortices. To connect this finding with previously mentioned, we can say that the subcortical branch holds raw information about the input from the environment and higher, neocortical branch holds knowledge of possible nuances or meta-information about the input. Saying differently, the higher in the hierarchy is the region, the more abstract information it processes. As Haker, Schneebeli, & Stephan (2016) pointed out when the abstract information is more precisely represented in the brain, we have a more certain interpretation of new experience, and at the same time, our internal model is sheltered from overfitting. In other words, we do not change entire internal model just because of extreme and surprising input from the environment.

The goal of this machinery is to be in the balance with the internal and external environment. Free-energy principle and predictive coding are theories of optimization (Friston, 2010, Friston, Kilner, & Harrison, 2006). More cognitively said, they describe how the system responds to the mismatch between predicted and perceived to remain homeostatic (Haker, Schneebeli, & Stephan, 2016).

2.2 Learning in Autism

To comprehend the whole picture of autism in the context of Bayesian Brain hypothesis is important to describe specificity of learning in autism. ASD individuals seem to have a problem with establishing intuition; therefore, implicit learning can be diminished. Collective of authors (Brown, Aczel, Jiménez, Kaufman & Grant, 2010) carried out a study of implicit learning in autism. Utilizing five implicit learning tasks in distinct cognitive domains (e.g. language, spatial-vision), they found no difference between ASD group and neurotypical control. There also was no correlation between task performance and autism symptomatology. These results are supported by similar experimental studies (e.g. Nemeth et al., 2010) as well as a meta-analysis of empirical research (Foti, Crescenzo, Vivanti, Menghini & Vicari, 2015). Hence, the poor implicit acquisition should be explained by another concept. FMRI study (Solomon et al., 2015) found that ASD individuals engage less cognitive control in transitive learning and therefore may have less integrated flexible learning. Authors, based on findings, claims, that it can normalize with age. Findings of Harris et al. (2015, 2016) show over-specificity in learning of ASD individuals and thus support previously mentioned. Summing up these outcomes, we can say that generalization is damaged.

Concerning generalization, there are two types present in high-functional ASD (Church et al., 2015). One type (T1) is like developmentally typical control and second (T2) is exact opposite. Type T1 generalize better when having more prototypes to learn from. Type T2 performs better when learning from one prototype and its detailed aspects. Authors didn't find any behavioral differences between groups probably because of small sample.
In general, learning of autistic people is typically spontaneous and sometimes characterized with extraordinary handling of complex material. They apparently are resistant to conventional education and have different implicit learning strategies (Dawson, Mottron & Gernsbacher, 2008).

2.3 Current Research

With the understanding of Bayesian Brain and learning in autism, we can continue with describing research in autism within the Bayesian framework. Haker, Schneebeli & Stephan (2016) describe ASD in a "Bayesian Brain" perspective through clinical symptoms - 1) perception in ASD is defined by weakly established abstract representations which provide predictions with low precision and fail in directing attention toward informative

stimuli; 2) repetitive behavior is caused by the previous point (i.e. it is easier to stay in predictable environment where abstract representations are not necessary); 3) social interaction is impaired due to high complexity of social context in which predictions of internal model have to be precise. For more details see Fig. 1.

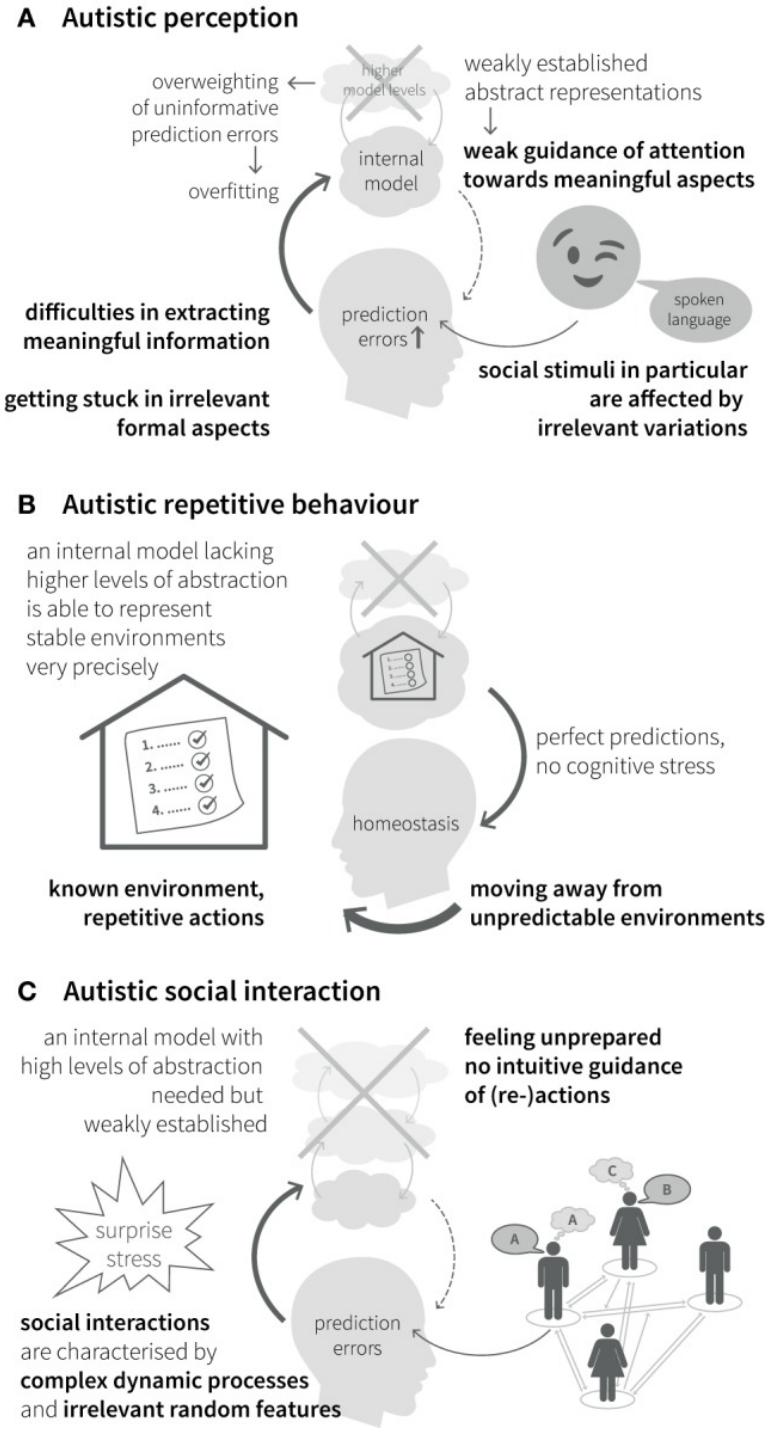


Fig. 1 Autistic brain from Bayesian perspective (Haker, Schneebeli, & Stephan, 2016, p. 8).

The often-cited article of Pellicano & Burr (2012) presents hypothesis of attenuated Bayesian priors (hypo-priors) as a root of autistic perception. Unprecise predictions of the internal model lead to the impression of higher accuracy of perceived input. In other words, autistic individuals are less affected by previous knowledge. Consequently, they notice much more details about the world. A Recent study of Pell et al. (2016) is first attempt to test this hypothesis within high-level of social perception. There are two important assumptions: 1) autistic people count on actual sensory input rather than on internal model of the world, and 2) neurotypical individuals tend to expect that other person's gaze is direct (i.e. direct gaze bias). In the study, authors investigated whether priors for direct gaze are reduced in uncertain condition comparing ASD and age and IQ-matched control group. They used visual social perception task where participants guess direction of the gaze of eyes from the picture. Additionally, they used visual noise to make the image of eyes blurry to simulate uncertainty. Sensitivity was measured with performing the task without the noise. Results support Bayesian Brain hypothesis because the increase of direct gaze bias was related to the increase of uncertainty (viewing eyes through visual noise). However, findings showed no difference between ASD and control group neither indirect gaze bias, nor in sensitivity. Although this may seem like an issue for Bayesian perspective on ASD, lack of differences can be explained by research limitations. The clinical sample consisted of high-functioning autistic individuals with mean age 32, not to mention a small number of participants. Empirical evidence from longitudinal study lasting almost five years (Shattuck et al., 2007) suggested that 31 and older ASD people have less maladaptive behavior and more often experience improvement. Besides, in 2016 similar study was published by Sevgi, Diaconescu, Tittgemeyer & Schilbach encouraging different results. In this case, computational analyses demonstrated difficulties in using social cues while making decisions within the unstable environment for participants with the greater level of autistic traits (Sevgi, Diaconescu, Tittgemeyer & Schilbach, 2016). Reduced sensitivity to context is diminished even in areas like motor tasks (Palmer, Paton, Kirkowski, Enticott & Hohwy, 2015). As we have already argued in previous part, ability to generalize is reduced in autism (Harris et al., 2015, Harris et al., 2016). In combination with preference of objects and lack of social motivation (Chevalier, Kohls, Troiani, Brodkin & Schultz, 2012) Bayesian account on autism seems reasonable. All previously stated theories and hypothesis fit together within this perspective. Bayesian Brain hypothesis can be perceived as an extension of them (Friston, Lawson & Frith, 2013).

3 Subtypes of Autism

Historically, the oldest subtypes of autism were explained by Kanner (1968) and Asperger (1944). These authors wrote case studies about autism independently. Kanner (1968) described children with infantile autism in this manner: Autistic children are usually recommended for screening for suspected severe mental retardation or hearing loss. Through investigation, it is found that their cognitive potential is covered with basal disabilities and eliminate hearing loss. The primary symptom is the inability to behave in a normal manner to the people. Almost all mothers reported the child's failure to respond to the gesture when parents indicated that they want to take the hand. Two-thirds of children can speak their language; however, there is a lack of exchange of information even in the cases with sufficient vocabulary where children remembered complicated words, poems, etc. Repeating words or phrases heard previously is present. Improper use pronouns - talking about himself in the second or third person. They marked by an obsessive-compulsive desire for maintaining an environment. With the remarkable accuracy to remember the layout of objects, toys, furniture and adversely react to small changes. They rather prefer objects than people. Often, there are some stereotypical physical manifestations. Although it seemed that the children are mentally disabled, with all present were indisputably in good cognitive condition.

Kanner's description is a very same description of one type of childhood schizophrenia described by Despert (1938). Asperger syndrome was typically portrayed as a less severe form of autism with milder developmental issues and better outcomes as age progresses. Still, the description seems to be almost the same (Wing, 1981): improper use of pronouns, repetitive speech, misinterpretation or ignorance of non-verbal communication, intensive attachment to possessions, clumsy and uncoordinated gross motor movements, etc. In empirical comparison to high-functioning autism, the difference is only in the level of severity (Ozonoff, South & Miller, 2000). Lack of differences supports the decision of APA to remove diagnosis of Asperger syndrome from diagnostic manual DSM-5 (APA, 2013).

To best of our knowledge, the first serious attempt to create a classification of autism came from Lorna Wing. Wing & Gould (1979) carried out screening of autism in one part of London. According to this pioneering research, autism has been diagnosed through the triad of symptoms for decades. Authors divided participant into four types considering social interactions. 1) *Aloof type* with little eye contact and faces empty of expressions, they ignore people and seem to have their "own world", 2) *Passive type* with some eye contact and

interaction but just passive involvement, 3) *Active but odd type* with eye contact and physical interaction are present but may be too strong or too long. They do not pay attention to others.

This nosology is supported by the research as well, although Beglinger & Smith (2001) highlighted that there are differences in the distribution of subtypes. Sometimes, there is the biggest cluster of aloof, at times, active-but-odd is most prevalent. Furthermore, in this review of up to date studies on autism subtyping, they suggested dimensional conceptualization (Fig.2) inspired by the conclusions from research studies. Fig. 2 shows that the most important features were social functioning, stereotyped behavior, and developmental delays. The number of possible groups on the continuum is at least four. However, the conceptualization is based on studies with different methodological problems. Some of them used datasets with mixed autistic children with genetically determined syndromes where autism is secondary.

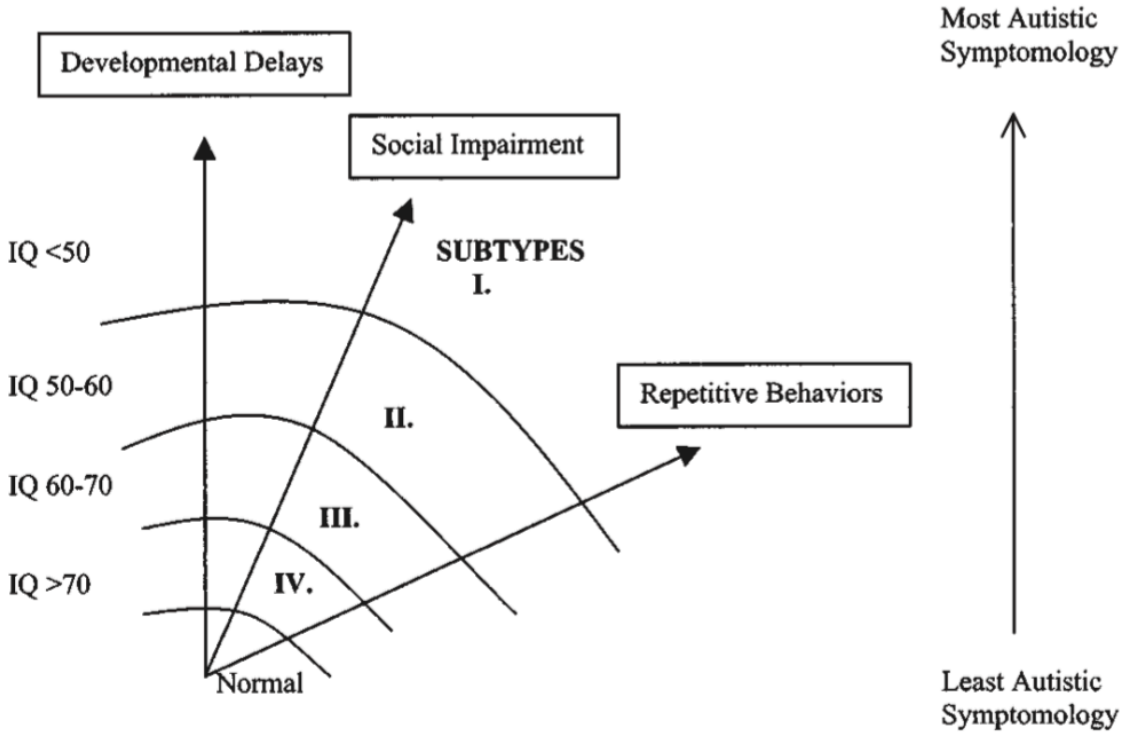


Fig. 2 Autism subtypes based on social deficits (Belinger & Smith, 2001, p. 420)

One of a hundred of autistic children has comorbid genetic conditions as Fragile-X syndrome, Rett syndrome, etc. (Caglayan, 2010). Therefore, this type of autism is also called syndromic. While studies considering comorbidities as subtypes of ASD contributed to the elucidation of core genetic pathologies (Caglayan, 2010), we should stay aware of the fact that we do not know how much syndromic and "pure" idiopathic autism differ. Idiopathic

autism is more common. Therefore, we will mainly focus on the behavioral analysis of such cases rather than analyze both groups as one.

Current studies in autism subtyping used comorbidity medical records (Doshi-Velez, Ge, & Kohane, 2014) or neuroimaging data (Hrdlička et al., 2005). Furthermore, behavioral data in combination with biological data (Veatch et al., 2014) or behavioral data only (Hu & Steinberg 2009; Shen, Lee, Holden & Shatkay, 2007; Eaves, Ho, & Eaves, 1994). Most often used clustering method is Agglomerative Hierarchical clustering (Shen, Lee, Holden & Shatkay, 2007; Hu & Steinberg 2009; Doshi-Velez, Ge, & Kohane, 2014; Veatch et al., 2014).

We picked up three studies which we consider important since they investigated behavioral subtypes for a detailed description. Shen, Lee, Holden & Shatkay (2007) performed three clustering methods and consensus clustering on behavioral assessment from ADI-R (Autism Diagnostic Interview - Revised). They included 358 participants in analysis and defined four distinct clusters. Two clusters were characterized with high severity slightly differing in language skills. However, both were considered nonverbal. Intermediate severity defined another cluster and last one was the least severe one across all domains and had late onset of symptoms. Hu & Steinberg (2009) obtained similar results with bigger data set – 1954 autistic individuals using hierarchical clustering only. A subset of individuals with syndromic autism was excluded. Results suggested subtypes based on severity, specifically mild, intermediate and very severe within the triad of symptoms. The last cluster had high savant skills. In comparison, Shen, Lee, Holden & Shatkay (2007) didn't find differences between their clusters in this domain. These findings also agree with the review on autism subtyping (Beglinger & Smith, 2001). However, Veatch et al. (2014) identify two main types and ten subtypes. The data set consisted of 1261 individuals with idiopathic ASD diagnosed with ADOS (Autism Diagnostic Observation Schedule), in addition to ADI-R. Such inclusion criteria are stricter in comparison to studies using one measurement. Two clusters were portrayed as "less severe" further splitting into four clusters and "more severe" further splitting into six clusters. Most important features were 1) onset of developmental abnormalities and 2) verbalism. ASD individuals in the "less severe" cluster were typically verbal and had later onset of symptoms. These clusters were genetically meaningful as well.

We wish to discuss subtypes of autism in the Bayesian framework; therefore, we sum up, sensory subtypes as well. Sensory subtypes are based on tests and questionnaires concerning sensory information processing. Such studies usually do not use behavioral data and therefore can fail to prove clinical relevance. The study of Lane, Molloy & Bishop

(2014) is one of the examples. Although they identified four sensory subtypes, there was no clinical difference between clusters. On the other hand, Ausderau et al. (2014) with more complex analysis, bigger sample, and one-year longitudinal design achieved better results. They found four sensory subtypes reflecting gender, age, the severity of symptoms and family characteristics. 1) The mild subtype with low autistic symptomatology and no sensory limitations/enhancements at all. 2) The sensitive-distressed cluster was typical with the highest hyperresponsiveness to sensory stimuli or oversensitivity with enhanced perception and highest intelligence. 3) The attenuated-preoccupied group had the lowest intelligence, strong autistic traits and were hyposensitive with sensory interests. 4) The extreme-mixed subtype had highest autistic traits and a high score in hyper- and hypo- responsiveness as well as enhanced perception and sensory interests. In the next exploration of these subtypes authors (Ausderau et al., 2016) discovered lowest ability to adapt in attenuated-preoccupied subtype and highest maladaptive behavior in extreme-mixed type. Classification of sensory subtypes by Audserau and colleagues is the best one up to this date according to a systematic review (DeBoth & Reynolds, 2017). DeBoth & Reynolds, however, emphasized the necessity of further investigation. Authors also pointed out the need for consensus and collaboration. Studies usually use very different measurements, and results are therefore incomparable.

Another critical view is that clusters can be a function of the age (Siegel, Anders, Ciaranello, Bienenstock, & Kraemer, 1986). Interpretation of such result can be in advantage of a spectral understanding of autism without subtypes. For instance, Mayes & Calhoun (2011) found a correlation of symptoms severity and age and intelligence. The severity of autism symptoms increased with decreasing chronological age and intelligence. Nevertheless, the study on animal models (Walsh, Morrow, Rubenstein, 2008) suggested that developmental window to form ASD symptoms may be wider than assumed. Specific defects can be related to synaptic activity and change of activity. Such reasoning suggests that autism can appear due to environmental factors after birth, what would explain why first symptoms show later in early development.

To connect possible subtypes of autism and Bayesian Brain hypothesis we can only synthesize previously mentioned. The authors sometimes note the possibility of Bayesian subtypes (e.g. Haker, Schneebeli & Stephan, 2016) but to best of our knowledge, nobody has ever taken the idea further. Considering the dimensional model of Belinger & Smith (2001) and sensory subtypes (Ausderau et al., 2016) there might be two main subtypes defining worst and normal but odd cases – 1) over-fitters and 2) special learners. The most

severe cases with bigger developmental delay may be over-fitters often consumed by sensory input. The problem in this group can be that they do not use internal model at all. For example, even if they have a stereotyped schedule in the school, they can have problems to follow it on time because “interesting” object steals their attention. These individuals can have sub-normal intelligence and verbal language skills, but with no interest in the social world, they are not motivated to have concepts in a first place. In sensory subtype terminology (Audserau et al., 2016), these are attenuated-preoccupied and extreme-mixed subtypes. On the other hand, the less severe group struggle with accurate weights for the internal model. Specific learners can use the internal model, and they could have established weights or concepts in specific, well-known situations. However, updating of the internal model weights or metacognition is slow and requires different strategies. Regarding sensory subtypes, these are sensitive-distressed and mild. Slower social learners could have a later onset.

4 Methods

4.1 Research Problem and Research Questions

Autism is considered as a spectrum. The question what causes the symptomatology described in previous chapters remains unanswered. Autism is associated with cognitive theories focusing on "local" way of thinking like Extreme Male Brain theory encouraging the role of fetal testosterone (Baron-Cohen, 2002) or Weak Central Coherence hypothesis (Happé, 1999). Extreme Male Brain theory as lack of Theory of Mind (Baron-Cohen, 2002) and Weak Central Coherence hypothesis (Happé, 1999) speculate about autism as about cognitive style (i.e. a way of thinking) because of its spectral nature. The theories describe an autistic cognitive style as very detail-focused and systematic. Concerning this approach, autistic individuals better comprehend local relationships associated with one object or system rather than global context. Describing autistic reasoning explains just cognitive symptomatology. Furthermore, the cognitive theories are related to neurobiological hypotheses suggesting anomalies in the amygdala (e.g. Baron-Cohen et al., 2000) or brain volume (e.g. Fidler, Bailey & Smalley, 2000). These concepts, however, have not provided a complex explanation of impairments in autism and cannot explicate possible subtypes. Quite recently, considerable attention has been paid to Bayesian Brain hypothesis which should be the one box for all previously mentioned views. Bayesian brain hypothesis (Knill, & Pouget, 2004) is a theory which understands the brain as a probabilistic machine which uses its general model to generate predictions. Current sensory input is tested against predictions or priors, and general internal model updates its beliefs about causes of the input. Pellicano & Burr (2012) proposed precision of the general model predictions to be attenuated (hypo-priors) in autism. The reason that Bayesian perspective is promising account for autism subtyping is a possibility to computationally model subtypes using various parameters. Computational modeling could help to understand the nature of the differences within autistic group better and explain inconsistencies in biological research. Additionally, Bayesian Brain hypothesis approach symptomatology in an interactive manner and therefore explain nature of autism better (Haker, Schneebeli & Stephan, 2016).

Former research in autism subtyping suggested at least four clusters (Beglinger, & Smith, 2001). Current research has conversely shown (Hu & Sternberg, 2009; Veatch et al., 2014) that, there are two main subtypes - "less severe" and "more severe". The two subtypes define extremes of a continuum regarding severity. Clusters between the extremes vary from research to research. However, last research investigating behavioral subtypes (Veatch et al.,

2014) in the large sample, identified two clusters with sub-clusters. The most important factors were an onset of symptoms and language delay; therefore, it will be interesting to test whether subtypes also differ from individuals with syndromic (secondary) autism and language deficit. From the Bayesian perspective, the "less severe" group could have slower social learning and "more severe" cluster could over-fit current sensory input. Such groups would have different requirements to approach social learning. In practice, this can affect therapies and create new methods for each type.

As noted before, popular theories of autism fail to explain empirical findings. The main problem is replication due to conflicting results. Inconsistencies may be explained by the existence of distinct types of autism with qualitatively same problems differing in quantity. A key limitation of autism subtyping research is that analyses are based on various individuals, usually diagnosed just with one diagnostic method. The poor diagnosis could cause inclusion of syndromic autistic children in which autism is a secondary consequence of genetic syndrome and misdiagnosed children with language deficits, ADHD, etc.. The research questions are formulated as follows:

Q1 What is the number of behavioral subtypes of autism?

Q2 Do non-autistic participants differ in behavioral description from autistic ones?

4.2 Research Sample

We received data from Autism Research Center at Comenius University, Faculty of Medicine. Although the original dataset included 313 individuals, we kept 281 for analysis mostly due to missing scores. The final sample consists of 217 autistic children, particularly twenty-nine girls (13.36%) and one hundred eighty-eight boys (86.64%). Further, nine girls (21.43%) and thirty-three boys (78.57%) having autistic traits as a comorbidity to a genetic syndrome or other developmental disorder; and seven girls (33.33%) and fourteen boys (66.66%) with language disorder. Detailed information about age is provided in Tab. 1.

Tab. 1 Descriptive statistics: Age across all groups

Group	N	Minimum	Maximum	Mean	SD
ASD: girls	29	2.1	20.6	7.2	4.64
ASD: boys	188	2.7	24.6	7.5	4.44
Secondary Autism: girls	9	3.0	14.5	7.5	4.58
Secondary Autism: boys	33	4.2	13.7	8.9	2.81
Language Disorder: girls	7	3.8	7.9	6.2	1.24
Language Disorder: boys	14	3.0	9.5	5.4	1.83

4.3 Measurements

The Autism Diagnostic Observational Schedule (ADOS) is one of the most commonly used observational methods together with its extended version (ADOS-2). ADOS-2 is the second, revised version of ADOS and it includes a new Module for Toddler, based on which autism can be diagnosed very early (Luyster et al., 2009). Early diagnosis is crucial for early intervention. However, we do not use forth module in our analysis. The three remaining modules consider three subtypes based on verbal ability and chronological age. This diagnostic scale focuses on the evaluation of communication, social interaction, and stereotyped behaviors. The protocol consists of a set of structured and partially structured interviews and activities involving the interaction between the clinician and the investigated subject. The administrator will get a picture of the three most important areas – 1) the language and communication skills, 2) mutual interaction, 3) games, stereotypical expressions and narrowly defined interests. All these areas get a single score. The most important are scores in communication and mutual social interaction. If the sum of the scores from both areas reaches a certain threshold, an individual is diagnosed with an autism spectrum disorder (Lord et al., 2000).

Activities provide children with an interesting, motivating and standardized context for social interactions through precisely defined games and tools that create the right environment for social interaction. Administering the ADOS-2 scale takes approximately 30-60 minutes. During this time, the administrator provides the child with a series of opportunities to uncover his social and communication abilities relevant to determining the diagnosis of autism (Kubranská, Vidošovičová, Kvasničková & Ostatníková, 2015).

Besides ADOS-2, Autism Diagnostic Interview - Revised (ADI-R) is one of the most proven methods based on the semi-structured interview with the parents of a child or adult with suspected symptoms of autism (Le Couteur, Haden, Hammal, & McConachie, 2008). The test is most reliable when used in the pre-school age. The biggest disadvantage is a long time of administration (about 90-120 minutes), which reduces the possibility of routine and widespread use in clinical practice. Scores obtained in ADI-R algorithm can be converted to ICD-10 criteria and gain four score key for diagnosis in the following areas (Lord, Rutter, & Le Couteur, 1994): 1) *Social Interaction:* Non-verbal behavior is evaluated as a mean of social interaction, as well as the ability to make contact with peers, ability to mutually share the pleasure and the level of socio-emotional skills. 2) *Communication:* In non-verbal communication, a capacity to use gestures is assessed and a symbolic and social level of

playing of mimic games. The goal in verbal communication is to evaluate the ability of conversation and peculiarities in speech development. 3) *Behavior with recurrent and stereotyped tendencies*: Evaluating movement patterns, narrowness and adhesion behavior in the interests of peace compulsive behavior, dysfunctional handling of objects and unusual sensory interests. 4) *Age, when the first symptoms occur and when the language is acquired*. All the items were scored by experienced clinicians from The Centre for Autism Research within Medical Faculty of Comenius University following instructions from ADI-R. The summary of scoring from Lord, Rutter, & Le Couteur (1994) is presented in Tab. 2. It is important to add that items are formulated as specific behavior, not as a question (e.g. social smiling or imaginative play with peers). The specific behavior is described in the manual, so clinicians always use the same criteria for behavior assessment. The number of items in total is 93. However, we used just 34 items because the rest is related to comorbidity and was not significant in the previous clustering.

Tab. 2 Scoring of ADI-R items

0	No definite behavior of the type specified
1	Behavior of the type specified probably present but defining criteria not fully met
2	Definite abnormal behavior of the type described in the definition and coding
3	Extreme severity (occasionally used)

4.4 Data Analysis

In this part of thesis, we illustrate the outline of analyses in a way that all methods are briefly described as well as reason of using them.

4.4.1 Pre-processing

Both ADI-R and ADOS subscales vary. By variation here we mean that not all items must be filled in due to verbal abilities and chronological age. The diagnostic algorithm lowers the score of extreme cases to objectivize and generalize an overall score.

In the first step, data were normalized. The normalization of the subscales points oscillates according to 1) the number of items in the subscale (from 2 to 5, the maximum number of points from 6 to 15); 2) Age criterion (e.g. items only for 3 or 4 years old); 3) verbal ability of the child; 4) how many items were really filled; 5) the maximum value of the item (predominantly 0-3 but also 0-2). Even if it sounds complicated, such normalization is quite straightforward. We take the individual maximum that the child could have reached

instead of using the highest score occurring in the dataset. Normalized scores range from zero to one as regularly. ADI-R contains two subscales involving only verbal children. Since we also wanted to use these subscales, we have given the non-verbal children the worst possible results as if they had a maximum of symptoms. Such imputation is recommended by Hu & Steinberg (2009).

In items for repetitive behavior, one of two items is selected, and only that one is included in the diagnostic score. The chosen one is the one with higher occurrence in child's behavior. Due to the possible redundancy of information, we followed this diagnostic trend in the pre-processing of the data.

Unfortunately, we obtained one of the ADI-R parts as categorical variables. This part (D) is holding information about the onset of abnormalities and onset of language. This subscale was quite significant in previous research (Veatch et al., 2014) therefore we made these items continuous with multiple correspondence analysis from FactoMineR R package (Lê, Josse, & Husson, 2008). The method is usually used to reduce dimensions of the data, and it creates new dimensions based on variance (Husson, Lê & Pagès, 2010). If we want to reduce dimensionality, we can use just principal components with the highest variance explained. However, this is not our case, and we used all the components. Thus, we have changed the type of the variable from nominal to cardinal.

4.4.2 Cluster Validation

Since our data analysis is exploratory, we consider necessary to evaluate different clustering solutions with cluster validation techniques and pick up the right number of clusters. Tan, Steinbach & Kumar (2006) classified cluster validity measures into three categories: 1) internal, 2) external, 3) relative. In the thesis, we use Average Silhouette width as internal index and Jaccard index as an external/relative measure. Jaccard index is formerly external measure, but we will use it differently, therefore in our case, it is a relative measure. We will describe details later.

Average Silhouette width is a measurement of fitness cluster fitness designed for K-means clustering (Rousseeuw, 1987). Thanks to this method we can decrease a level of subjectivity in decision making concerning the number of clusters. The main idea behind the average silhouette width is a comparison of within-cluster with between-cluster dissimilarities for every participant and then average across the cluster or all the clusters (Kodinariya & Makwana, 2013; Rousseeuw, 1987). The greater the difference between

those two distances, the better the result. Silhouette width ranges from zero to one, where one means perfectly compact and well-separated clusters.

Jaccard coefficient is one of the mostly used cluster validity measures (Tan, Steinbach & Kumar, 2006). It is a similarity-oriented index (Tan, Steinbach & Kumar, 2006), indicating the extent of intersection (Wagner & Wagner, 2007) of two clustering results ranging from zero to one. One shows identical clustering result. Its disadvantage is sensitivity to cluster size and number of clusters. In our case, we used the coefficient as a comparison measure for hierarchical clustering of our dataset versus hierarchical clustering on permutated dataset. Final Jaccard index is averaged across 1000 random permutations.

4.4.3 Cluster Analysis

As we already mentioned before the most often used method in autism subtyping is hierarchical clustering (HLC). HLC permit clusters to have sub-clusters and can be perceived as a sequence of partition in hierarchical order (i.e. next partition is related to previous one). It is especially useful in finding small clusters. We derive centers of clusters (cluster prototypes), and then we use these prototypes as initiation points for model-based density clustering. Density-based clustering methods are recommended when analyzing noisy data with outliers. Clinical sample is such dataset. (Tan, Steinbach & Kumar, 2006)

Agglomerative Hierarchical Clustering "start with the points as individual clusters and at each step, merge the closest pair of clusters. This requires notion of cluster proximity" (Tan, Steinbach & Kumar, 2006, p.515). In our case, proximity was defined with Ward's method which represents clusters as centroids. According to Blashfield's experiments (1976), Ward's method is the most accurate and stable. Stability is important criterion since we test stability with Jaccard index. Concerning software, we used factoextra R package (Kassambara, & Mundt, 2015) to perform agglomerative hierarchical clustering and visualize dendrogram and individual silhouette width plot.

To perform *Model-based clustering*, we used mclust R package (Scrucca, Fop, Murphy & Raftery, 2016). The package was intended for model-based clustering, classification, and density estimation based on finite normal mixture modeling. It implements functions for parameter estimation via the Expectation-Maximization algorithm (EM) for normal mixture models with a variety of covariance structures. Moreover, it offers functions for simulation from these models and combines EM for mixture estimation with the Bayesian Information Criterion (BIC) in a whole strategy for clustering and density estimation. Explaining how exactly the model-based clustering works is beyond the scope

of the thesis. What we wish to emphasize is that this way of clustering offers very stable solutions. Interpretation will be more specific (see chapter 5 Results).

4.4.4 Analysis after Clustering

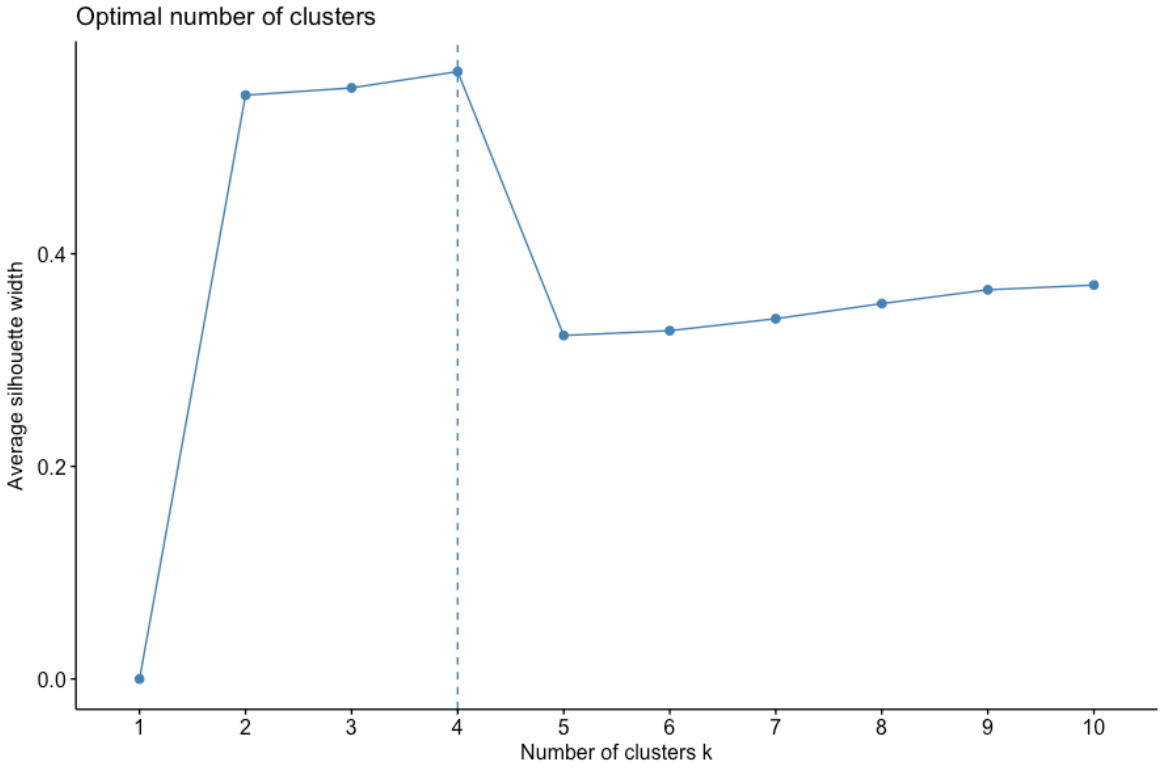
In this exploratory step, we will identify crucial variable for clustering results and compare autistic clusters with non-autistic individuals. We particularly plan to use ETA squared as a measure of association for a combination of nominal and cardinal variables. In our case, it is the amount of variation in behavioral score explained by clusters. Then, we will use Pearson chi-squared test and Cramer's V to investigate an association between clusters and nominal variables concerning developmental issues of autistic individuals.

In the end, we will compare clusters and non-autistic sub-sample with Student t-test for independent samples or its non-parametric version Wilcoxon Test for independent samples (mostly known as Mann-Whitney U test). We will decide according to results of normality tests for terminal groups. All tests mentioned in this part are implemented in R bases, except Pearson Chi-squared from Coin R package (Horthorn, Hornik, Wiel & Zeileis, 2008).

5 Results

Results are presented about research questions. First research question *Q1* was formulated as follows: *What is the number of behavioral subtypes of autism?* As described in analysis description, we computed an optimal number of clusters with Averaged Silhouette width for the number of clusters ranging from two to ten. The method suggested four clusters using hierarchical cluster analysis utilizing Euclidian distance and Ward's linkage method. As shown in Fig. 3, the most optimal number of clusters is four. As can be additionally seen, the three and the two cluster solutions are optimal comparably. In all three alternatives, the silhouette would be more than 0.50 what indicate compact clusters. We also measured the stability of the clusters by comparing permutated and original dataset utilizing Jaccard Index. Coefficients suggest very similar results as Averaged Silhouette width. The similarity of two clustering results drops with five clusters. For more information, the reader is referred to Tab. 3.

Fig. 3 Average Silhouette width: Optimal number of clusters for HLC



Tab. 3 Jaccard index: HLC on permutated and original dataset

N of Clusters	Jaccard Index	N of Clusters	Jaccard Index
2	**0.5270**	**7**	0.2796
3	**0.5075**	**8**	0.2769
4	**0.5029**	**9**	0.2765
5	0.2817	**10**	0.2763
6	0.2815		

We are open to possibility of two complementary clustering results, therefore, we took four clusters over two or three clusters as the best solution. Fig. 4 depict dendrogram of hierarchical tree with one huge cluster (N(C1) = 175) and three small ones (N(C2) = 20; N(C2) = 15; N(C4) = 7). We identified one misplaced individual with Silhouette width plot for individual cases as presented in Fig. 5.

Fig. 4 Dendrogram: HCL, four cluster solution

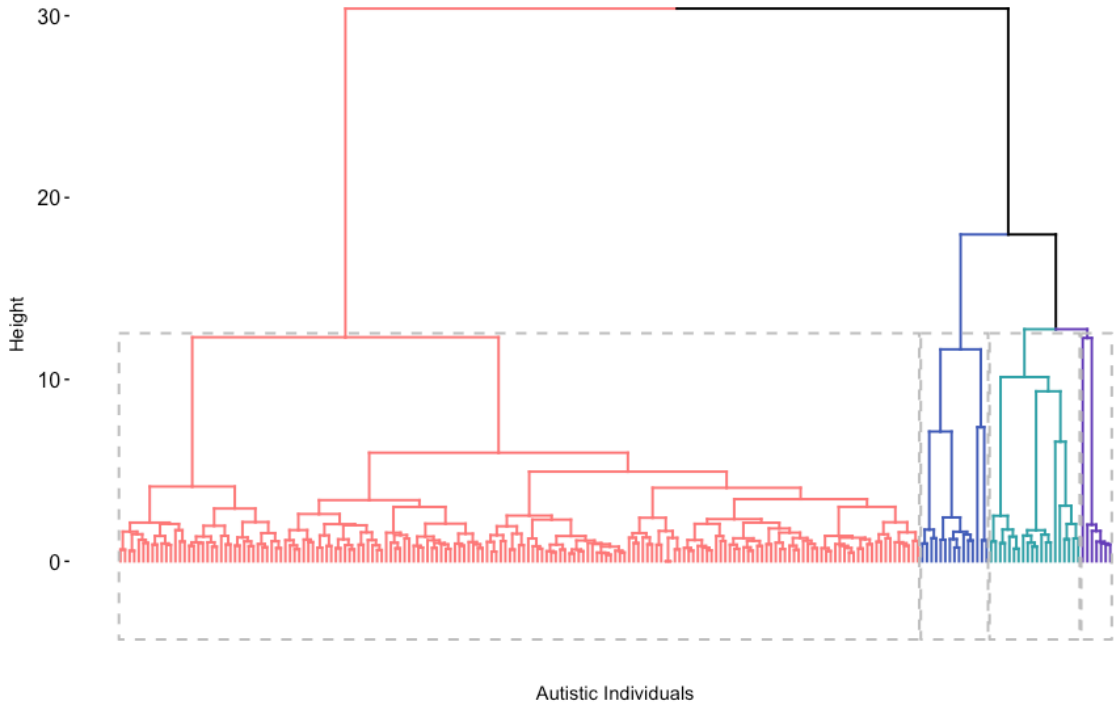
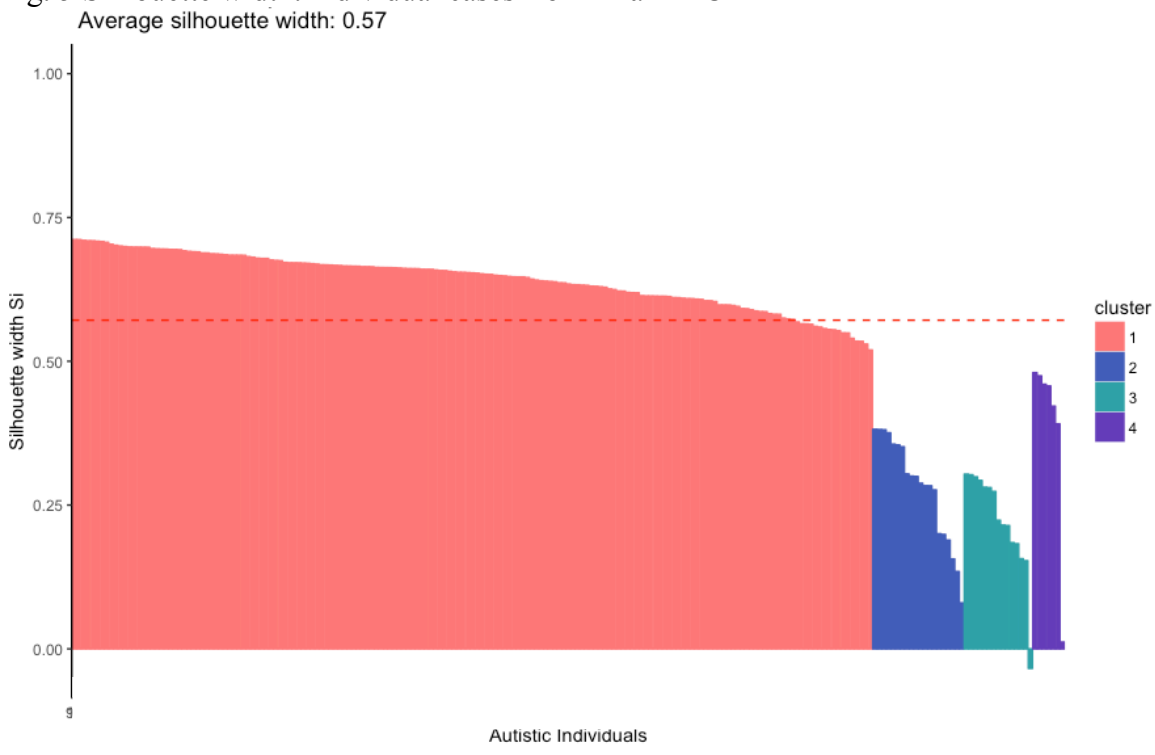
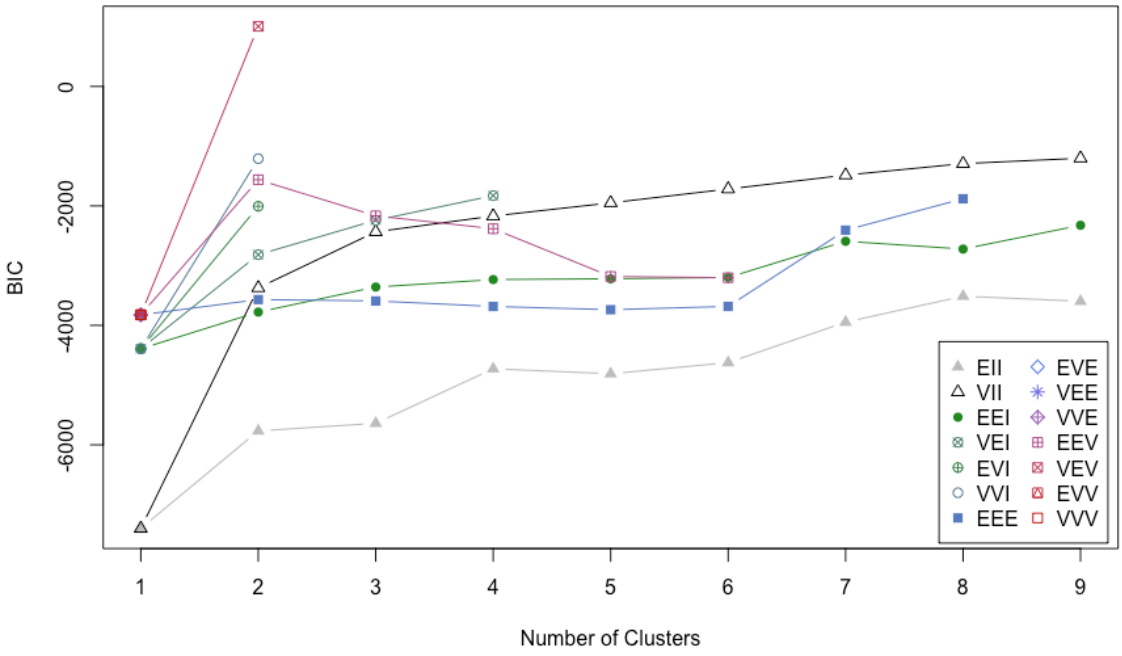


Fig. 5 Silhouette width: individual cases from final HLC



In the next step, we used centers of clusters from Hierarchical clustering as initialization points for model-based clustering. Primarily, we computed Bayesian Information Criterion (BIC) for different numbers of clusters and models. The higher the BIC the better. Fig. 6 outlines the best number of clusters is two. The result is compatible with findings from our previous analysis. The best model is abbreviated "VEV" indicating that clusters are ellipsoidal and have equal shape.

Fig. 6 Bayesian Information Criterion: Optimal number of clusters

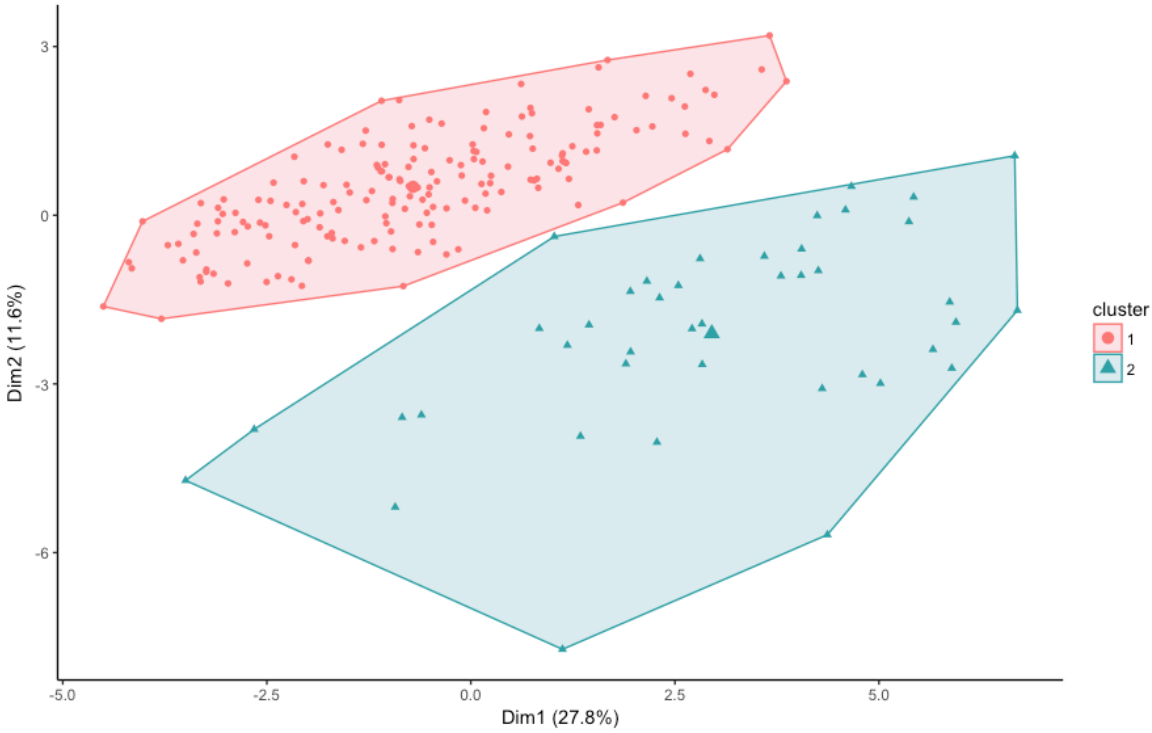


Furthermore, we used final model with highest BIC for model-based cluster analysis. We obtained two distinct clusters where one is nicely shaped, and second one is opposite and does not seem compact. Mixing probabilities are low as well (see Tab. 4). Low mixing probabilities indicate that individuals in the second cluster does not form one cluster. In other words, these cases can be misplaced. Tab. 4 summarizes model parameters.

Tab. 4 Model Parameters

Log.likelihood	**N**	**df**	**BIC**	**Mixing Prob. Cluster 1**	**Mixing Prob. Cluster 2**
1005.9	217	485	1005.9	0.81	0.19

Fig. 7 Cluster plot: model-based clustering, two clusters



Final step within first research question is to explore how clusters where formed. Thus, we want to know to what extend is a variation of ADOS and ADI-R subscales significantly explained by clusters. According to ETA-squared index, we can tell that Gesture Communication and stereotyped speech in ADI-R had medium effect on clustering result (see Tab. 5). Other ADI-R subscales – Interaction Regulation, Peer Relationships, Odd Interests and Ritualistic Behavior, are not included in the Tab. 5 due to no significance. It is necessary to emphasize that subscales Stereotyped Speech and Conversational Interchange are originally meant only for verbal children, and non-verbal children had imputed the worst possible score. We will compare clusters in these domains later.

Tab. 5 ETA-squared: Variation explained by clusters

	ETA squared	p
ADI-R Gesture Communication	0.13	8.41E-08
ADI-R Stereotyped Speech (V)	0.12	2.52E-07
ADI-R Conversational Interchange (V)	0.10	1.17E-06
ADI-R Sensory Issues	0.09	8.40E-06
ADOS Stereotyped Behavior and Interests	0.07	7.02E-05
ADI-R Stereotyped Motor Mannerism	0.07	7.91E-05
ADI-R Socio-emotional Reciprocity	0.06	3.45E-04
ADI-R Social Development and Game	0.06	4.32E-04
ADOS Social Reciprocity	0.05	1.50E-03
ADI-R Shared Enjoyment	0.04	2.06E-03
ADOS Communication	0.02	2.98E-02

We tested association between questions concerning developmental onset and final clusters. These are the questions from ADI-R, which were pre-processed to cardinal variable. To see which question was important we used Pearson Chi-squared test and Cramer's V. We will describe every question separately.

Tab. 6 shows cross-tabulation of clusters and question regarding caregiver (parents) estimate of age of abnormalities onset. Chi-squared test indicated that the question and clustering results are dependent (X=111.6, p = 0.000) and association is strong (V = 0.72). As can be seen from Tab. 6, in cluster 1 was only one individual whom symptoms had shown later than in 36 months of his age.

Tab. 6 Pearson Chi squared test – Parent estimate of abnormalities by clusters

	Parent estimate of abnormalities		**Chi2**	**p**	**Cramer's V**
	> 36 months	**<36 months**			
Cluster 1	1	174	111.6	**0.000**	**0.72**
Cluster 2	25	17			

Tab. 7 shows Pearson chi-squared test, Cramer's V test and cross-tabulation of child's estimated age of first words and clusters. Results suggest independence of variables and Cramer's V indicates no relationship.

Tab. 7 Pearson Chi squared test – Parent estimate of first words by clusters

	Parent estimate of first words		**Chi2**	**p**	**Cramer's V**
	< 24 months	**>24 months**			
Cluster 1	75	100	1.9	0.164	0.10
Cluster 2	23	19			

Tab. 8 presents comparison of final clusters and question regarding first sentences. Results of Chi-squared test and Cramer's V test shows that variables are dependent and there is medium association. Thus, cluster 2 had been more-or-less typically developing and most of the individuals in first cluster had delayed onset of first sentences.

Tab. 8 Pearson Chi squared test – Parent estimate of first sentences by clusters

	Parent estimate of first sentences		**Chi2**	**p**	**Cramer's V**
	< 33 months	**>33 months**			
Cluster 1	37	138	40.1	**0.000**	**0.43**
Cluster 2	30	12			

Tab. 9 shows again comparison of question about onset of abnormalities and final clusters. However, in this instance, the estimation is done by questionnaire itself. In other words, the answer to this question depends on previous three answers. We identified dependency of clusters and the question (X = 30.1; p = 0.000) with medium association (V = 0.37). In the first cluster are only individuals with early onset (Tab. 9).

Tab. 9 Pearson Chi squared test – Questionnaire estimate of an onset by clusters

	Questionnaire estimate of onset		**Chi2**	**p**	**Cramer's V**
	Late onset	**Early onset**			
Cluster 1	0	175	30.1	**0.000**	**0.37**
Cluster 2	22	20			

The final comparison of question with clusters is presented in Tab. 10. We can conclude that variables are dependent to each other (X = 102.0; p = 0.000) and have strong association (V = 0.69). Clinicians view is in line with estimate of a parent or other caregiver, and have strongest association with clustering results.

Tab. 10 Pearson Chi squared test – Administrator estimate of abnormalities by clusters

	Administrator estimate of abnormalities		**Chi2**	**p**	**Cramer's V**
	> 36 months	**<36 months**			
Cluster 1	0	175	102.0	**0.000**	**0.69**
Cluster 2	7	35			

In comparison of two clusters with Wilcoxon Test for independent samples we found differences across all subscales except Interaction Regulation, Peer Relationships, Odd Interests, Ritualistic Behavior from ADI-R and Communication from ADOS. Biggest effect size was identified in case of Gesture Communication (p = 0.000; r = 0.35). Cluster 2 has generally better results, meaning less severe symptoms (Tab. 11).

Tab. 11 Wilcoxon W test - cluster 1 vs. cluster 2

variable	group	N	Median	mean rank	W	sig.	r
A1	Cluster 1	175	0.44	112.4	3980.5	0.099	-
	Cluster 2	42	0.33	94.8			
A2	Cluster 1	175	0.67	111.8	4088	0.178	-
	Cluster 2	42	0.58	97.3			
A3	Cluster 1	175	0.63	114.9	3548.5	**0.004**	**0.21**
	Cluster 2	42	0.44	84.5			
A4	Cluster 1	175	0.50	116.2	3326.5	**0.001**	**0.24**
	Cluster 2	42	0.33	79.2			
B1	Cluster 1	175	0.67	119.3	2780.5	**0.000**	**0.35**
	Cluster 2	42	0.33	66.2			
B4	Cluster 1	175	0.67	115.5	3434.5	**0.002**	**0.24**
	Cluster 2	42	0.44	81.8			
B2	Cluster 1	175	1.0	118.7	2878	**0.000**	**0.32**
	Cluster 2	42	0.6	68.5			
B3	Cluster 1	175	1.0	119.3	2778	**0.000**	**0.34**
	Cluster 2	42	0.39	66.1			
C1	Cluster 1	175	0.33	105.13	18397.5	0.059	-
	Cluster 2	42	0.33	125.1			
C2	Cluster 1	175	0.33	108.3	18945.5	0.714	-
	Cluster 2	42	0.33	112.1			
C3	Cluster 1	175	0.67	116.4	3286	**0.000**	**0.27**
	Cluster 2	42	0.33	78.2			
C4	Cluster 1	175	0.67	115.9	3375.5	**0.001**	**0.29**
	Cluster 2	42	0.50	80.4			
COM	Cluster 1	175	0.60	113.4	3817.5	0.037	-
	Cluster 2	42	0.50	90.9			
RSI	Cluster 1	175	0.65	115.4	3454	**0.002**	**0.21**
	Cluster 2	42	0.51	82.2			
UVRS	Cluster 1	175	0.42	117.3	3121	**0.000**	**0.27**
	Cluster 2	42	0.25	74.3			

*Note 1: Holm-Bonfferoni Correction for p-value

*Note 2: A1 – Interaction Regulation; A2 – Peer Relationship, A3 – Shared Enjoyment; A4 – Socio-emotional Reciprocity; B1 – Gesture Communication; B4 – Social Development and Game, B2 – Conversational Interchange, B3 – Stereotyped Speech; C1 – Odd Interests; C2 – Ritualistic Behavior; C3 – Stereotyped Motor Mannerism; C4 – Sensory Issues; COM – Communication; RSI – Social Reciprocity; UVRS – Stereotyped Behavior and Interests

In the second part of the results, we answer second research question *Q2* which was formulated as follows: *Do non-autistic participants differ in behavioral description from autistic ones?* Results show that second cluster is generally more similar to non-autistic group than first cluster.

In case of cluster one and syndromic autistic group we found differences in Conversational Interchange ($p = 0.000$), Shared Enjoyment ($p = 0.000$), Social Development and Game, Communication and Social Reciprocity with strong effect ranging from $r = 0.50$ to $r = 0.60$. On the other hand, scores in Ritualistic Behavior are not statistically significant at all and scores in Odd Interest sub-scale are significantly different but with small effect size ($r = 0.19$). When comparing first cluster to individuals with language deficits we found comparable results. No significant differences in Ritualistic Behavior and differences with small effect size are in Odd Interest ($p = 0.001$; $r = 0.21$), Stereotyped Speech ($p = 0.007$; $r = 0.21$), Stereotyped Motor Mannerism ($p = 0.001$; $r = 0.24$) and Interaction Regulation ($p = 0.000$; $r = 0.28$). Overall effect sizes from this comparison are smaller than from comparison of cluster one and syndromic autism. Detailed results can be found in Tab. 12 (comparison with syndromic autism) and Tab. 13 (comparison with language disorder).

Comparing cluster two with syndromic autism (Tab. 14) and language disorders (Tab. 15) brings less differences than first cluster with non-autistic individuals. Besides, we found differences with strong effect size in Communication and Social Reciprocity in both comparisons ranging from $r = 0.51$ to $r = 0.56$. We found no differences between cluster 2 and non-autistic groups in Socio-Emotional Reciprocity, Gesture Communication, Conversational Interchange, Stereotyped Speech, Ritualistic Behavior and Stereotyped Behavior and Interest. Group with language deficit do not differ in Interaction Regulation as well. Other between-groups differences have medium effect size.

Tab. 12 Wilcoxon W test - cluster 1 vs. secondary autism

variable	group	N	Median	mean rank	W	sig.	r
A1	Cluster 1	175	0.44	121.6	2589.5	**0.000**	**0.38**
	Secon. a.	43	0.11	60.22			
A2	Cluster 1	175	0.67	121.4	2635.5	**0.000**	**0.39**
	Secon. a.	43	0.33	61.3			
A3	Cluster 1	175	0.63	124.7	2049	**0.000**	**0.53**
	Secon. a.	43	0.0	47.7			
A4	Cluster 1	175	0.50	123.8	2200	**0.000**	**0.46**
	Secon. a.	43	0.21	51.2			
B1	Cluster 1	175	0.67	123.8	2199	**0.000**	**0.49**
	Secon. a.	43	0.11	51.1			
B4	Cluster 1	175	0.67	125	2002.5	**0.000**	**0.56**
	Secon. a.	43	0.11	46.6			
B2	Cluster 1	175	1.0	121.7	2567.5	**0.000**	**0.50**
	Secon. a.	43	0.40	59.7			
B3	Cluster 1	175	1.0	119.7	2929.5	**0.000**	**0.35**
	Secon. a.	43	0.3	68.1			
C1	Cluster 1	175	0.33	115.1	3732.5	**0.007**	**0.19**
	Secon. a.	43	0.17	86.7			
C2	Cluster 1	175	0.33	112.5	4184	0.142	-
	Secon. a.	43	0.17	97.3			
C3	Cluster 1	175	0.67	119.8	2905	**0.000**	**0.34**
	Secon. a.	43	0.0	67.6			
C4	Cluster 1	175	0.67	119.1	3036.5	**0.000**	**0.35**
	Secon. a.	43	0.50	70.6			
COM	Cluster 1	175	0.60	124.7	1608	**0.000**	**0.54**
	Secon. a.	43	0.11	39.2			
RSI	Cluster 1	175	0.65	124.6	1631.5	**0.000**	**0.60**
	Secon. a.	43	0.16	39.8			
UVRS	Cluster 1	175	0.42	120.8	2301	**0.000**	**0.37**
	Secon. a.	43	0.08	56.1			

*Note 1: Holm-Bonfferoni Correction for p-value

*Note 2: A1 – Interaction Regulation; A2 – Peer Relationship, A3 – Shared Enjoyment; A4 – Socio-emotional Reciprocity; B1 – Gesture Communication; B4 – Social Development and Game, B2 – Conversational Interchange, B3 – Stereotyped Speech; C1 – Odd Interests; C2 – Ritualistic Behavior; C3 – Stereotyped Motor Mannerism; C4 – Sensory Issues; COM – Communication; RSI – Social Reciprocity; UVRS – Stereotyped Behavior and Interests

Tab. 13 Wilcoxon W test - cluster 1 vs. language disorder

variable	group	N	Median	mean rank	W	sig.	r
A1	Cluster 1	175	0.44	103.3	1033.5	**0.000**	**0.28**
	lang. dys.	20	0.11	51.7			
A2	Cluster 1	175	0.67	105.6	634	**0.000**	**0.45**
	lang. dys.	20	0.17	31.7			
A3	Cluster 1	175	0.63	105.75	604	**0.000**	**0.46**
	lang. dys.	20	0.07	30.2			
A4	Cluster 1	175	0.50	105.9	585	**0.000**	**0.43**
	lang. dys.	20	0.14	29.3			
B1	Cluster 1	175	0.67	104.4	846.5	**0.000**	**0.37**
	lang. dys.	20	0.11	42.3			
B4	Cluster 1	175	0.67	106.3	506	**0.000**	**0.50**
	lang. dys.	20	0.11	25.3			
B2	Cluster 1	175	1.0	102.5	1177	**0.000**	**0.38**
	lang. dys.	20	0.4	58.9			
B3	Cluster 1	175	1.0	101.4	1368.5	**0.007**	**0.21**
	lang. dys.	20	0.40	68.4			
C1	Cluster 1	175	0.33	102.5	1172.5	**0.001**	**0.21**
	lang. dys.	20	0.0	58.6			
C2	Cluster 1	175	0.33	100.4	1737	0.160	-
	lang. dys.	20	0.17	82.7			
C3	Cluster 1	175	0.67	103.1	1268.5	**0.001**	**0.24**
	lang. dys.	20	0.33	60.4			
C4	Cluster 1	175	0.67	104.4	849.5	**0.000**	**0.39**
	lang. dys.	20	0.33	42.5			
COM	Cluster 1	175	0.60	106.4	689.5	**0.000**	**0.43**
	lang. dys.	21	0.13	32.8			
RSI	Cluster 1	175	0.65	107.7	458	**0.000**	**0.54**
	lang. dys.	21	0.20	21.8			
UVRS	Cluster 1	175	0.42	105.4	863	**0.000**	**0.32**
	lang. dys.	21	0.17	41.1			

*Note 1: Holm-Bonfferoni Correction for p-value

*Note 2: A1 – Interaction Regulation; A2 – Peer Relationship, A3 – Shared Enjoyment; A4 – Socio-emotional Reciprocity; B1 – Gesture Communication; B4 – Social Development and Game, B2 – Conversational Interchange, B3 – Stereotyped Speech; C1 – Odd Interests; C2 – Ritualistic Behavior; C3 – Stereotyped Motor Mannerism; C4 – Sensory Issues; COM – Communication; RSI – Social Reciprocity; UVRS – Stereotyped Behavior and Interests

Tab. 14 Wilcoxon W test - cluster 2 vs. secondary autism

variable	group	N	Median	mean rank	W	sig.	r
A1	Cluster 2	42	0.33	52.6	1446.5	**0.000**	**0.36**
	Secon. a.	43	0.11	33.6			
A2	Cluster 2	42	0.58	52.2	1462.5	**0.001**	**0.36**
	Secon. a.	43	0.33	34.0			
A3	Cluster 2	42	0.44	54.2	1379.5	**0.000**	**0.42**
	Secon. a.	43	0.0	32.1			
A4	Cluster 2	42	0.33	49.5	1575	0.015	-
	Secon. a.	43	0.21	36.6			
B1	Cluster 2	42	0.33	48.8	1604.5	0.028	-
	Secon. a.	43	0.11	37.3			
B4	Cluster 2	42	0.44	52.8	1437	**0.000**	**0.41**
	Secon. a.	43	0.11	33.4			
B2	Cluster 2	42	0.60	49.4	1579	0.015	-
	Secon. a.	43	0.40	36.7			
B3	Cluster 2	42	0.39	44.2	1800	0.664	-
	Secon. a.	43	0.30	41.9			
C1	Cluster 2	42	0.33	52.6	1446.5	**0.000**	**0.41**
	Secon. a.	43	0.17	33.6			
C2	Cluster 2	42	0.33	46.5	1701.5	0.177	-
	Secon. a.	43	0.17	39.7			
C3	Cluster 2	42	0.33	46.0	1723	0.222	-
	Secon. a.	43	0.0	40.1			
C4	Cluster 2	42	0.50	45.5	1743	0.336	-
	Secon. a.	43	0.50	40.5			
COM	Cluster 2	42	0.50	55.8	1143	**0.000**	**0.56**
	Secon. a.	43	0.11	27.9			
RSI	Cluster 2	42	0.51	55.2	1169	**0.000**	**0.54**
	Secon. a.	43	0.16	28.5			
UVRS	Cluster 2	42	0.25	48.4	1452	0.013	-
	Secon. a.	43	0.08	35.4			

*Note 1: Holm-Bonfferoni Correction for p-value

*Note 2: A1 – Interaction Regulation; A2 – Peer Relationship, A3 – Shared Enjoyment; A4 – Socio-emotional Reciprocity; B1 – Gesture Communication; B4 – Social Development and Game, B2 – Conversational Interchange, B3 – Stereotyped Speech; C1 – Odd Interests; C2 – Ritualistic Behavior; C3 – Stereotyped Motor Mannerism; C4 – Sensory Issues; COM – Communication; RSI – Social Reciprocity; UVRS – Stereotyped Behavior and Interests

Tab. 15 Wilcoxon W test - cluster 2 vs. language disorder

variable	group	N	Median	mean rank	W	sig.	r
A1	Cluster 2	42	0.33	35.5	462	0.010	**-**
	lang. dys.	20	0.11	23.1			
A2	Cluster 2	42	0.58	38.2	350.5	**0.000**	**0.54**
	lang. dys.	20	0.17	17.5			
A3	Cluster 2	42	0.44	37.2	391.5	**0.000**	**0.45**
	lang. dys.	20	0.07	19.6			
A4	Cluster 2	42	0.33	35.4	465.5	0.130	-
	lang. dys.	20	0.14	23.3			
B1	Cluster 2	42	0.33	33.2	559.5	0.279	-
	lang. dys.	20	0.11	28.0			
B4	Cluster 2	42	0.44	36.9	404.5	**0.001**	**0.44**
	lang. dys.	20	0.11	20.2			
B2	Cluster 2	42	0.60	33.4	549	0.213	-
	lang. dys.	20	0.40	27.5			
B3	Cluster 2	42	0.39	30.5	1282.5	0.538	-
	lang. dys.	20	0.40	33.5			
C1	Cluster 2	42	0.33	37.3	388.5	**0.000**	**0.45**
	lang. dys.	20	0.0	19.4			
C2	Cluster 2	42	0.33	34.4	572.5	0.132	-
	lang. dys.	20	0.17	27.3			
C3	Cluster 2	42	0.33	32.7	642	0.638	-
	lang. dys.	20	0.33	30.6			
C4	Cluster 2	42	0.50	34.6	500	0.043	-
	lang. dys.	20	0.33	25.0			
COM	Cluster 2	42	0.50	38.5	400	**0.000**	**0.51**
	lang. dys.	21	0.13	19.1			
RSI	Cluster 2	42	0.51	39.4	361	**0.000**	**0.54**
	lang. dys.	21	0.20	17.2			
UVRS	Cluster 2	42	0.25	35.5	526	0.032	-
	lang. dys.	21	0.17	25.1			

*Note 1: Holm-Bonfferoni Correction for p-value

*Note 2: A1 – Interaction Regulation; A2 – Peer Relationship, A3 – Shared Enjoyment; A4 – Socio-emotional Reciprocity; B1 – Gesture Communication; B4 – Social Development and Game, B2 – Conversational Interchange, B3 – Stereotyped Speech; C1 – Odd Interests; C2 – Ritualistic Behavior; C3 – Stereotyped Motor Mannerism; C4 – Sensory Issues; COM – Communication; RSI – Social Reciprocity; UVRS – Stereotyped Behavior and Interests

6 Discussion

Regarding the scope of the thesis, we identified two statistically different subtypes of autism. These subtypes differed from non-autistic but behaviorally similar groups. The results indicate an existence of two clusters of autistic individuals. One group consists of three-fourths of the whole sample and the second group includes remaining individuals. The second cluster could be interpreted as a group of outliers with greater similarity to the non-autistic group. Our findings have more than one interpretation in the context of theories.

An important question in autism research is whether autism is a spectrum of different disorders or it is one disorder with spectral symptomatology (Lai, Lombardo, Charrabarti & Baron-Cohen, 2014). This question is difficult to answer, and research is not clear; however, clinicians diagnosing children every day need to have a concept of the disorder. This fact is related to DSM V. (APA, 2013) where autism is just one diagnosis – Autism Spectrum Disorder, with different specifications based on symptoms. In the context of our results and previous research, we support the spectral view due to practical clinical issues. There can be even fifteen subtypes, but at this point, research does not have particular answers. As Constantino (2011) pointed out traditional categorical diagnostic approaches may lead to misclassification of subjects - mostly girls and mildly affected boys in multiple-incidence autism families. According to the author, the misclassification can be particularly damaging to biological studies. Although we agree with the spectral nature of autism, we disagree with categorical approach damaging biological studies. The main reason is that something practical for clinicians does not have to come in handy for researchers. Biologically speaking, autism is the consequence of various pathophysiologic pathways (Chaste & Leboyer, 2012) and it is very likely there are subtypes due to conflicting results. The autism research results are not confusing because of excluded or never included participants, but those who are included and should not be.

Back to our data, we found one quite compact group and then the group of outliers. The group of autistic participants consisted of individuals diagnosed with ADOS and with ADI-R as well. It was necessary to pass the certain threshold on both measurements, and despite strict criteria, we found forty-two outliers divided into three small sub-clusters. Something similar was observed in the study of Veatch et al. (2014). Our final dendrograms were considerably similar on first sight. Such subtypes structures may suggest that there are

subtypes which are unequally sized. Miles et al. (2005) found two clusters as well. Authors name them complex autism and essential autism. A complex type is characterized by the great percentage of syndromic autistic individuals and significant brain dysmorphology and microcephaly. In the essential cluster, there were no comorbidities, and this cluster contained eighty percent of the whole sample. Our results are in line with these findings because our second cluster was more similarities with non-autistic groups than the first cluster. The second cluster had less severe autistic symptoms, and this leads to another issue. We suggest considering a possibility that mild autism can have higher variation in symptoms as the more serious one. Such view is supported by the study of Hahamy, Behrmann & Malach (2015) proposing that individualized changes in functional brain connectivity organization be a core characteristic of high-functioning autism, and can be the solution for contradictory findings. They also did not found clear subtypes of ASD in their dataset suggesting either no subtypes or subtle groups which do not show in the small dataset.

The largest difference occurred in the onset of developmental anomalies. The most likely explanation is that two core biological roads of autism exist. One with early onset and more severe symptoms and the second one with later onset and high variation in severity. The reason for this may be a various amount of environmental and genetic factors contributing to the condition. From the Bayesian perspective, this would mean that less severe group is more affected by the environment because an internal model of these individuals depends on their everyday life. Therefore, there is more variability. This line of thinking is consistent with social motivation theory. Social motivation theory reflects on possible subtypes based on diminished social motivation (Chevalier, Kohls, Troiani, Brodkin & Schultz, 2012). Authors hypothesize that active-but-odd subtype has social motivation. Weak Central Coherence and Bayesian Brain hypotheses do not explain social domain of autistic symptomatology, therefore struggle with explaining subtypes alone. However, a combination of the cognitive approach of Extreme Male Brain theory, and perceptual approach of the Bayesian brain with motivational factors explain results quite well. Severe autistic individuals prefer closed systems rather than uncertain open systems, they are focused on unimportant details, and perceptual experiences are still interesting even after thousands of repetitions. They are pleased enough from the low-level sensory information only, and they do not need social interaction. Social domain is usually avoided in theories probably because of difficulties of the operationalization of variables. However, research of Chang et al. (2014) suggests different reasoning. Authors found that connectivity of social-emotional pathways of autistic individuals is different from people with sensory

processing disorders. They sensory processing pathways, though, are similar. These results suggest that sensory problems are not sufficient for other autistic symptomatology and this confront pure cognitive theories.

Our results also show the importance of gesture communication domain. Verbal and nonverbal individuals formed mixed groups; therefore, it may suggest that motivation to communicate is more important than the ability to speak. In terms of neurotransmitters, oxytocin is responsible for social memory and behavior and therefore is promising therapeutic agent of autism-specific symptoms. Few randomized controlled trials have already been successful (Yoo, 2015). Concerning Bayesian Brain hypothesis is oxytocin perceived as a modulator of precision or weights of prediction error (Friston, 2016). In other words, malfunction of oxytocin functionality and low accuracy of internal model predictions are related.

Concerning qualitative findings, autistic groups were different in all ADI-R and ADOS domains, except Interaction Regulation, Peer Relationships, Odd Interests and Repetitive Behaviors. All exceptions are domains which child could compensate with age and more experiences. Nonetheless, we found mostly small effect sizes. We identified medium effect size in 1) gesture communication, which we have already discussed, and 2) verbal domains, which probably result from the greater number of nonverbal individuals in first, bigger cluster. We also found that ADOS performs better in distinguishing between non-autistic and autistic groups, especially Communication and Social Reciprocity subscales. We identified largest effect sizes.

6.1 Limitations

One of the technical limitations is that determining the number of clusters is without significance testing and we decide just according to indices. Furthermore, when we tested differences between autistic and non-autistic groups, the group with language disability was smaller than the statistical minimum. We also found suspicious that all the groups have shown no differences in scores of repetitive behaviors. There is a chance that this is false negative significance level and we did not found differences which were present.

The data-related limitation is that our sample is quite artificial. Data were collected at the University in specialized workplace where parents go for a first time to diagnose a child. We also missed biological data which were incomplete at the date of analysis but will be part of the future analysis.

All the limitations restricted possibility to draw strong or final conclusions; therefore, further investigation is needed.

6.2 Future Directions

Future research should combine behavioral data with sensory subtypes and biological data. We would like to emphasize the importance of longitudinal studies because age plays probably important role. Research should focus on early development, labor, and pregnancy to find possible biomarkers. It seems like behavioral subtypes may be confusing especially with mixed diagnoses and small research samples. On the other hand, to carry out screening of autism in newborns or embryos is quite difficult in practice. Mostly due to ethical issues. Thus, the first step towards more consistent results can be the extraction of behavioral outliers before analysis of biological data. This point is again connected to the number of participants. Researchers are usually happy even for a small sample and remove some cases can be very hurtful to a final number of analyzed cases. Fortunately, current databases and collaborative research make the problem with sample size easier and easier.

Another likely way can be linking genes to particular behavioral symptoms as St. Pourcain (2013). Authors investigated 6948 children from the general population and found two new genetic correlates of social communication, which have never been mentioned in relation to ASD. Subtypes could be falsified with investigating biological roots of early onset ASD group in comparison with the late-onset group.

Conclusion

The thesis aimed to discuss subtypes of autism in the context of current theories and analyze behavioral data. In the theoretical part, we described autism, actual problems, and present theories. Then, we focused on Bayesian Brain hypothesis and Autism in Bayesian Framework. The last theoretical chapter characterized subtypes of autism from the historical perspective as well as the point of view of current research. Our research problem and questions were, and then we described empirical methods and data analysis pipeline. Theories of autism fail to explain empirical findings because of conflicting results and weak explanatory power. The existence of subtypes would explain conflicting results and consensus of opinions would offer better explanations. A fundamental limitation of autism subtyping research is that analyses are based on various individuals, usually diagnosed just with one diagnostic method. Such approach could cause inclusion of syndromic autistic children in which autism is a secondary consequence of genetic syndrome and misdiagnosed children with language deficits, ADHD, etc. We attempted solving these issues in practical part of the thesis.

Apparently, one of the problems is theoretical and second one is a matter of approaching analysis. From the theoretical point of view, we concluded that despite Bayesian Brain theory offers robust explanations, it cannot account for an interpretation of broad autistic symptomatology. Therefore, we suggest connecting Bayesian framework with ideas of social motivation theory. Such combination can better explain the existence of subtypes as well. We found two clusters. The first cluster is bigger than second, and individuals are mostly nonverbal having greater scores in ADOS and ADI-R than in the second one. Therefore, the first one is “more severe”. The second cluster is smaller with verbal participants, and cases in this group are more similar to non-autistic individuals. The largest difference between clusters is an onset of developmental anomalies. All instances in the first cluster had early onset of symptoms. Regarding theoretical conclusion, the onset of symptoms can be related to a level of motivation to social stimuli. With a small degree of motivation to the social world, one does not need to form abstract concepts and representations. Abstract concepts and generalization are necessary for having precise predictions coming from the internal model. Such reasoning is in line with subtypes differing in ability to share enjoyment and interact reciprocally.

However, generalizations are limited due to described limitations. Despite this fact, we believe that the thesis contributed to the field of autism research with new ideas and useful recommendations.

References:

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.

Ausderau, K., Sideris, J., Furlong, M., Little, L. M., Bulluck, J., & Baranek, G. T. (2014). National survey of sensory features in children with ASD: Factor structure of the sensory experience questionnaire (3.0). *Journal of autism and developmental disorders*, 44(4), 915-925.

Ausderau, K. K., Sideris, J., Little, L. M., Furlong, M., Bulluck, J. C., & Baranek, G. T. (2016). Sensory subtypes and associated outcomes in children with autism spectrum disorders. *Autism Research*, 9(12), 1316-1327.

Aylward, E. H., Minshew, N. J., Goldstein, G., Honeycutt, N. A., Augustine, A. M., Yates, K. O., ... & Pearlson, G. D. (1999). MRI volumes of amygdala and hippocampus in non–mentally retarded autistic adolescents and adults. *Neurology*, 53(9), 2145-2145.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in cognitive sciences*, 6(6), 248-254.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience & Biobehavioral Reviews*, 24(3), 355-364.

Blashfield, R. K. (1976). Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. *Psychological Bulletin*, 83(3), 377.

Beglinger, L. J., & Smith, T. H. (2001). A review of subtyping in autism and proposed dimensional classification model. *Journal of Autism and Developmental Disorders*, 31(4), 411-422.

Brown, J., Aczel, B., Jiménez, L., Kaufman, S. B., & Grant, K. P. (2010). Intact implicit learning in autism spectrum conditions. *The quarterly journal of experimental psychology,* 63(9), 1789-1812.

Caglayan, A. O. (2010). Genetic causes of syndromic and non-syndromic autism. *Developmental Medicine & Child Neurology*, 52(2), 130-138.

Chang, Y. S., Owen, J. P., Desai, S. S., Hill, S. S., Arnett, A. B., Harris, J., ... & Mukherjee, P. (2014). Autism and sensory processing disorders: shared white matter disruption in sensory pathways but divergent connectivity in social-emotional pathways. *PloS one*, 9(7), e103038.

Chaste, P., & Leboyer, M. (2012). Autism risk factors: genes, environment, and gene-environment interactions. *Dialogues in Clinical Neuroscience*, 14(3), 281-92.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in cognitive sciences*, 16(4), 231-239.

Church, B. A., Rice, C. L., Dovgopoly, A., Lopata, C. J., Thomeer, M. L., Nelson, A., & Mercado, E. (2015). Learning, plasticity, and atypical generalization in children with autism. *Psychonomic bulletin & review*, 22(5), 1342-1348.

Constantino, J. N. (2011). The quantitative nature of autistic social impairment. *Pediatric research*, 69(5 Pt 2), 55R.

Courchesne, E., Campbell, K., & Solso, S. (2011). Brain growth across the life span in autism: age-specific changes in anatomical pathology. *Brain research*, 1380, 138-145.

DeBoth, K. K., & Reynolds, S. (2017). A systematic review of sensory-based autism subtypes. *Research in Autism Spectrum Disorders*, 36, 44-56.

Despert, L. J. (1938). Schizophrenia in children. *Psychiatric Quarterly*, 12(2), 366-371.

Dawson, M., Mottron, L., & Gernsbacher, M. A. (2008). Learning in autism. Learning and memory: A comprehensive reference, 2, 759-772.

Doshi-Velez, F., Ge, Y., & Kohane, I. (2014). Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. *Pediatrics*, *133*(1), e54-e63.

Eaves, L. C., Ho, H. H., & Eaves, D. M. (1994). Subtypes of autism by cluster analysis. *Journal of Autism and Developmental Disorders*, *24*(1), 3-22.

Fidler, D. J., Bailey, J. N., & Smalley, S. L. (2000). Macrocephaly in autism and other pervasive developmental disorders. *Developmental Medicine & Child Neurology*, 42(11), 737-740.

Foti, F., De Crescenzo, F., Vivanti, G., Menghini, D., & Vicari, S. (2015). Implicit learning in individuals with autism spectrum disorders: a meta-analysis. *Psychological Medicine*, 45(05), 897-910.

Freitag, C. M., Luders, E., Hulst, H. E., Narr, K. L., Thompson, P. M., Toga, A. W., ... & Konrad, C. (2009). Total brain volume and corpus callosum size in medication-naive adolescents and young adults with autism spectrum disorder. *Biological psychiatry*, 66(4), 316-319.

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature Reviews Neuroscience,* 11(2), 127-138.

Friston, K. (2016). The Bayesian Savant. *Biological Psychiatry*, 80(2), 87-89.

Friston, K. J., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: comment on Pellicano and Burr. *Trends in Cognitive Science*, 17(1), 10-1016.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, *100*(1), 70-87.

Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind". *Cognition*, 50(1), 115-132.

Grelotti, D. J., Gauthier, I., & Schultz, R. T. (2002). Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Developmental psychobiology*, 40(3), 213-225.

Gleiser, P. M., & Spoormaker, V. I. (2010). Modelling hierarchical structure in functional brain networks. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *368*(1933), 5633-5644.

Hahamy, A., Behrmann, M., & Malach, R. (2015). The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nature neuroscience*, *18*(2), 302-309.

Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian theories of autism spectrum disorder Help improve Clinical practice?. *Frontiers in psychiatry*, *7*.

Happé, F. (1999). Autism: cognitive deficit or cognitive style?. *Trends in cognitive sciences*, 3(6), 216-222.

Happé, F., Frith, U., & Briskman, J. (2001). Exploring the cognitive phenotype of autism: weak "central coherence" in parents and siblings of children with autism: I. Experimental tests. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(3), 299-307.

Harris, H., Israeli, D., Minshew, N., Bonneh, Y., Heeger, D. J., Behrmann, M., & Sagi, D. (2015). Perceptual learning in autism: over-specificity and possible remedies. *Nature neuroscience*.

Harris, H., Israeli, D., Minshew, N. J., Bonneh, Y. S., Heeger, D. J., Behrmann, M., & Sagi, D. (2016). Response: Commentary: Perceptual learning in autism: over-specificity and possible remedies. *Frontiers in Integrative Neuroscience*, 10.

Hothorn, T., Hornik, K., van de Wiel, M. A. and Zeileis, A. (2008). Implementing a class of permutation tests: the coin package. *Journal of Statistical Software* **28**(8), 1–23.http://www.jstatsoft.org/v28/i08/

Hrdlicka, M., Dudova, I., Beranova, I., Lisy, J., Belsan, T., Neuwirth, J., ... & Blatny, M. (2005). Subtypes of autism by cluster analysis based on structural MRI data. *European child & adolescent psychiatry*, 14(3), 138-144.

Hu, V. W., & Steinberg, M. E. (2009). Novel clustering of items from the Autism Diagnostic Interview-Revised to define phenotypes within autism spectrum disorders. *Autism Research*, *2*(2), 67-77.

Husson, F., Le, S. and Pages, J. (2010). Exploratory Multivariate Analysis by Example Using R, *Chapman and Hall*.

Kassambara, A., & Mundt, F. (2015). Factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0. 3.

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, *1*(6), 90-95.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in neurosciences*, *27*(12), 712-719.

Kubranská, A, Vidošovičová, M., Kvasničková, J., Ostatníková, D. (2015) Diagnostika porúch autistického spektra. *Psychiatria-psychoterapia-psychosomatika*, 22(1), p. 26-31

Lai, M. C., Lombardo, M. V., Chakrabarti, B., & Baron-Cohen, S. (2013). Subgrouping the Autism “Spectrum": Reflections on DSM-5. *PLoS Biol*, 11(4), e1001544.

Lane, A. E., Molloy, C. A., & Bishop, S. L. (2014). Classification of children with autism spectrum disorder by sensory subtype: A case for sensory-based phenotypes. *Autism Research*, 7(3), 322-333.

Larson, F. V., Lai, M. C., Wagner, A. P., Baron-Cohen, S., Holland, A. J., & MRC AIMS Consortium. (2015). Testing the ‘Extreme Female Brain’theory of psychosis in adults with autism spectrum disorder with or without co-morbid psychosis. *PloS one*, 10(6), e0128102.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of statistical software*, 25(1), 1-18.

Le Couteur, A., Haden, G., Hammal, D., & McConachie, H. (2008). Diagnosing autism spectrum disorders in pre-school children using two standardised assessment

instruments: the ADI-R and the ADOS. *Journal of autism and developmental disorders*, *38*(2), 362-372.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, 30(3), 205-223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders*, *24*(5), 659-685.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., ... & Richler, J. (2009). The Autism Diagnostic Observation Schedule—Toddler Module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of autism and developmental disorders*, 39(9), 1305-1320.

Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., ... & Deen, B. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, *19*(6), 659-667.

Mayes, S. D., & Calhoun, S. L. (2011). Impact of IQ, age, SES, gender, and race on autistic symptoms. *Research in Autism Spectrum Disorders*, 5(2), 749-757.

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in human neuroscience*, *5*, 39.

Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., ... & Farmer, J. E. (2005). Essential versus complex autism: definition of fundamental prognostic subtypes. *American Journal of Medical Genetics Part A*, 135(2), 171-180.

Müller, R. A., Shih, P., Keehn, B., Deyoe, J. R., Leyden, K. M., & Shukla, D. K. (2011). Underconnected, but how? A survey of functional connectivity MRI studies in autism spectrum disorders. *Cerebral Cortex*, *21*(10), 2233-2243.

Nemeth, D., & Janacsek, K. (2010). The dynamics of implicit skill consolidation in young and elderly adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbq063.

Ozonoff, S., South, M., & Miller, J. N. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from high-functioning autism. *Autism*, *4*(1), 29-46.

Palmen, S. J., Pol, H. E. H., Kemner, C., Schnack, H. G., Durston, S., Lahuis, B. E., ... & Van Engeland, H. (2005). Increased gray-matter volume in medication-naive high-functioning children with autism spectrum disorder. *Psychological medicine*, 35(04), 561-570.

Palmer, C. J., Paton, B., Kirkovski, M., Enticott, P. G., & Hohwy, J. (2015). Context sensitivity in action decreases along the autism spectrum: a predictive processing perspective. *Proceedings of the Royal Society of London B: Biological Sciences*, 282(1802), 20141557.

Pell, P. J., Mareschal, I., Calder, A. J., von dem Hagen, E. A., Clifford, C. W., Baron-Cohen, S., & Ewbank, M. P. (2016). Intact priors for gaze direction in adults with high-functioning autism spectrum conditions. *Molecular autism*, 7(1), 25.

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in cognitive sciences*, 16(10), 504-510.

Pellicano, L. (2012). Autism: Beyond weak central coherence.

Redcay, E., & Courchesne, E. (2005). When is the brain enlarged in autism? A meta-analysis of all brain size reports. *Biological psychiatry*, 58(1), 1-9.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20, 53-65.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *The R Journal*, 8(1), 289.

Sevgi, M., Diaconescu, A. O., Tittgemeyer, M., & Schilbach, L. (2016). Social Bayes: Using Bayesian modeling to study autistic trait–related differences in social cognition. *Biological psychiatry*, 80(2), 112-119.

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., ... & Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of autism and developmental disorders*, 37(9), 1735-1747.

Shen, J. J., Lee, P. H., Holden, J. J., & Shatkay, H. (2007). Using cluster ensemble and validation to identify subtypes of pervasive developmental disorders. In *AMIA Annual Symposium Proceedings* (Vol. 2007, p. 666). American Medical Informatics Association.

Siegel, B., Anders, T. F., Ciaranello, R. D., Bienenstock, B., & Kraemer, H. C. (1986). Empirically derived subclassification of the autistic syndrome. *Journal of Autism and Developmental Disorders*, 16(3), 275-293.

Solomon, M., Frank, M. J., Ragland, J. D., Smith, A. C., Niendam, T. A., Lesh, T. A., ... & Carter, C. S. (2015). Feedback-driven trial-by-trial learning in autism spectrum disorders. *American Journal of Psychiatry*, 172(2), 173-181.

St Pourcain, B., Whitehouse, A. O., Ang, W. Q., Warrington, N. M., Glessner, J. T., Wang, K., ... & McArdle, W. L. (2013). Common variation contributes to the genetic architecture of social communication traits. *Molecular autism*, 4(1), 34.

Swanson, L. R. (2016). The Predictive Processing Paradigm Has Roots in Kant. *Frontiers in systems neuroscience*, *10*.

Tan, P. N., Steinbach, M., & Kumar, V. (2006). Classification: basic concepts, decision trees, and model evaluation. Introduction to data mining, 1, 145-205.

Teatero, M. L., & Netley, C. (2013). A critical review of the research on the extreme male brain theory and digit ratio (2D: 4D). *Journal of autism and developmental disorders*, 43(11), 2664-2676.

Thorová, K. (2006). "Poruchy autistického spectra", 1st ed., Czech Republic, Prague: Portál

Vanegas, S. B., & Davidson, D. (2015). Investigating distinct and related contributions of weak central coherence, executive dysfunction, and systemizing theories to the cognitive profiles of children with autism spectrum disorders and typically developing children. Research in Autism Spectrum Disorders, 11, 77-92.

Veatch, O. J., Veenstra-VanderWeele, J., Potter, M., Pericak-Vance, M. A., & Haines, J. L. (2014). Genetically meaningful phenotypic subgroups in autism spectrum disorders. *Genes, Brain and Behavior*, 13(3), 276-285.

Wagner, S., & Wagner, D. (2007). Comparing clusterings: an overview. Karlsruhe: Universität Karlsruhe, Fakultät für Informatik.

Walsh, C. A., Morrow, E. M., & Rubenstein, J. L. (2008). Autism and brain development. *Cell*, 135(3), 396-400.

Whitehouse, A. J. (2016). Commentary: Are we expecting too much from the extreme male brain theory of autism? A reflection on Kung et al.(2016). *Journal of Child Psychology and Psychiatry*, 57(12), 1463-1464.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological medicine*, 11(01), 115-129.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of autism and developmental disorders*, 9(1), 11-29.

World Health Organization. (1993). The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research.

Yoo, H. (2015). Genetics of autism spectrum disorder: current status and possible clinical applications. *Experimental neurobiology*, 24(4), 257-272.

Zalla, T., & Sperduti, M. (2013). The amygdala and the relevance detection theory of autism: an evolutionary perspective. *Frontiers in human neuroscience*, 7, 894.

Zielinski, B. A., Prigge, M. B., Nielsen, J. A., Froehlich, A. L., Abildskov, T. J., Anderson, J. S., ... & Alexander, A. L. (2014). Longitudinal changes in cortical thickness in autism and typical development. *Brain*, 137(6), 1799-1812.