6-1-2013

Effectiveness of Gluten-Free and Casein-Free Diets for Individuals with Autism Spectrum Disorders: An Evidence-Based Research Synthesis

Jie Zhang
The College at Brockport, State University of New York

Michael R. Mayton
West Virginia University

John J. Wheeler
East Tennessee State University, wheelerjj@etsu.edu

Citation Information

Zhang, Jie; Mayton, Michael R.; and Wheeler, John J.. 2013. Effectiveness of Gluten-Free and Casein-Free Diets for Individuals with Autism Spectrum Disorders: An Evidence-Based Research Synthesis. *Education and Training in Autism and Developmental Disabilities*. Vol.48 276-286. http://daddcec.org/Portals/0/CEC/Autism_Disabilities/Research/Publications/Education_Training_Development_Disabilities/Full_Journals/ETADD_48_2_276-287.pdf ISSN: 2154-1647

Effectiveness of Gluten-Free and Casein-Free Diets for Individuals with Autism Spectrum Disorders: An Evidence-Based Research Synthesis

Education and Training in Autism and Developmental Disabilities, 2013, 48(2), 276–287

Effectiveness of Gluten-Free and Casein-Free Diets for Individuals with Autism Spectrum Disorders: An Evidence-Based Research Synthesis

Jie Zhang
The College at Brockport,
State University of New York

Michael R. Mayton
West Virginia University

John J. Wheeler
East Tennessee State University

Abstract: In order to better assist practitioners and better serve persons with autism spectrum disorders (ASD) and their families, it is vital for professionals to systematically evaluate the existing body of literature and synthesize its scientific evidence, so that the efficacy of research can be translated to evidence-based practices (EBPs) (Wheeler, 2007; Zhang & Wheeler, 2011). This research synthesis evaluated adherence to EBP standards and analyzed the effectiveness of gluten-free and casein-free (GFCF) diets for individuals with ASD. Four hundred and seventy articles were screened among peer-reviewed journals in English language published through 2010 using the Academic Search Complete search database. Twenty-three studies were selected, and the researchers used a systematic analysis model developed by Mayton, Wheeler, Menendez, and Zhang (2010) to investigate the degree of adherence to specific evidence-based practice standards. In addition, the study utilized quality indicators proposed by (a) Horner et al. (2005) for single-subject design studies and (b) Gersten et al. (2005) for group experimental design, to evaluate the efficacy of GFCF diet interventions. Results of this synthesis indicated that the efficacy of GFCF diet interventions for individuals with ASD is inconclusive, and the field needs better controlled studies to provide the scientific evidence base for the intervention.

Both the Elementary and Secondary Education Act of 2001 (No Child Left Behind, 2001) and the Individuals with Disabilities Education Improvement Act of 2004 (IDEiA, 2004) draw attention to the need for using scientifically-validated and evidence-based practices (EBPs). The Council for Exceptional Children (CEC, 2006) identified a research-based intervention as one that meets the following criteria: (a) four high quality studies with an effect on performance at .05 confidence level, (b) at least five single subject studies with adequate design and experimental control, or (c) studies conducted at three different settings by three different researchers.

Correspondence concerning this article should be addressed to Jie Zhang, Education and Human Development, The College at Brockport, State University of New York, 350 New Campus Drive, Brockport, New York 14420. E-mail: jzhang@brockport.edu

Given the dramatically increasing prevalence rates of individuals being diagnosed with autism spectrum disorders (ASD), it is urgent to conduct critical analyses of the extant research. The estimated prevalence rate of children with ASD has increased from 1 in 150 children (Centers for Disease Control and Prevention, 2007) to 1 in 110 children (CDC, 2010) in three years. Thus, it is important to identify and validate EBPs from a broad range of intervention practices to benefit individuals with ASD. In order to better assist practitioners and better serve people with ASD and their families, it is vital for professionals to systematically evaluate the existing body of literature and synthesize its scientific evidence, so that the efficacy of research can be translated to EBPs (Wheeler, 2007; Zhang & Wheeler, 2011).

Gluten-Free Casein-Free (GFCF) Diet Interventions for Individuals with ASD

Gluten-free casein-free interventions limit food that contains gluten (e.g. breads, pastas, pizza, bagels, crackers, cakes, cookies, oats/cereals, etc., made from wheat, barley, and rye) and casein (e.g., milk, cheese, cheese products, yogurt, ice creams, dips, sour cream, dressings, etc.). The existing studies on GFCF diet interventions are aimed at preventing gluten or casein from entering the bloodstream and thereby (theoretically) reducing/eliminating the symptoms of autism (Munasinghe, Oliff, Finn, & Wray, 2010). Christison and Ivany (2006) summarized four overlapping biological theories that support the GFCF diet interventions: opioid excess, reduced peptidase activity, immune dysfunction or autoimmunity, and gastrointestinal abnormalities. The opioid excess theory, the most popular theory supporting the GFCF diet interventions, hypothesized that abnormal leakage of gluten and casein related metabolites with opioid agonist properties from the gut pass into the central nervous system (CNS) and lead to intensified brain opioid activity and disrupted brain function (Christison & Ivany, 2006; Whiteley & Shattock, 2002).

Inevitably, the adoption of the GFCF diets may have some impact on families, including higher cost of special/alternative food products, extra time in food purchase/preparation, greater risk of nutrient deficiencies, and possible needs for additional supplements (e.g., calcium, vitamins, multivitamin with minerals). Families with individuals with ASD frequently learn about the treatment of GFCF diets and their efficacy as an intervention from the popular media. However, popular reports often do not address important questions such as: How effective is the intervention? Is this intervention supported by scientific evidence? Did all the existing studies report benefits and positive outcomes, or were there any adverse side effects? The purpose of this research synthesis was to evaluate adherence to EBP standards and analyze the effectiveness of the gluten-free and casein-free (GFCF) diet interventions for individuals with ASD.

Method

We examined studies published in peer-reviewed journals in English language between 1977 and 2010. To evaluate the efficacy of GFCF diet interventions, the researchers used a systematic analysis model developed from (a) the evidence-based standards and indicators developed by Mayton, Wheeler, Menendez, and Zhang (2010), (b) Horner et al. (2005) criteria for the evaluation of single-subject design studies, and (c) the Gersten et al. (2005) standards for the evaluation of group, experimental designs. Each selected study was evaluated across the following categories: participants' characteristics, characteristics of the dependent variables (DV), features of the independent variables (IV), and features of the study.

Criteria for Inclusion

Studies selected met three inclusion criteria, as follows: (a) the study used gluten-free and/or casein-free (GFCF) interventions; (b) the participants of the study included individuals diagnosed with ASD; and (c) all the articles were published in peer-reviewed journals in English language between 1977 and 2010.

Search Procedure

Searches were carried out to obtain articles for inclusion in this research synthesis. An electronic search was conducted using the *Academic Search Complete* database. Using one search term from each of the two categories, combinations of two keyword entries were used to select studies for the present synthesis: (a) autism, autistic, Asperger syndrome, Rett syndrome, Pervasive Developmental Disability-Not Otherwise Specified (PDD-NOS), autism spectrum disorders (ASD); and (b) gluten, casein, nutrition, diet. The total number of combined keyword sets was 24 (6x4). The search process resulted in the identification of 470 articles. After eliminating duplicates and excluding all irrelevant articles (e.g., reviews and position papers), 21 studies were retained. Additionally, relevant studies found in the reference section of the reviewed articles were located and chosen according to the se-

lection criteria. Both searches resulted in a total of 23 articles, which included 462 participants with ASD from 15 journals.

Data Coding

The researchers used a systematic analysis model developed by Mayton, Wheeler, Menendez, and Zhang (2010) to investigate the degree of adherence to evidence-based standards. In addition, the study incorporated quality indicators proposed by Horner et al. (2005) for single-subject design studies, as well as indicators from Gersten et al. (2005) for group, experimental design studies.

The authors used a coding sheet developed from the indicators identified within Mayton et al. (2010), Horner et al. (2005), and Gersten et al. (2005) during both the data coding and double-coding procedures. Using a systematic set of rules and procedures, the researchers coded relevant data from each selected study across the following categories: (a) participants' characteristics, including number of participants, age, gender, diagnosis, and the selection process; (b) characteristics of the dependent variables (DV), including the DV, operational definition, quantifiable measurement, valid and precise procedure description, maintenance and generalization, interrater reliability, and social validity; (c) features of the independent variables (IV), including the intervention, replicable description, systematic manipulation, and treatment integrity; and (d) features of the study, including research design, duration of the intervention, and results of the study. Data were transferred from hand-written coding sheets to an electronic spreadsheet, and then to SPSS for analysis.

Interrater Reliability

Interrater reliability was calculated by dividing the number of agreements by the total number of agreements and disagreements and multiplying by 100. The first two authors independently double-coded the data across 22 coding categories from all 23 articles ($N = 506$) and determined the interrater reliability. The mean interrater reliability was 97.4%, ranging from 78% to 100% within each category.

Analyses of Data

Descriptive data from the selected studies were analyzed by calculating both the frequency and percentage for each of the variables from the coding sheet. The percentages were calculated by dividing the number of items in a subset by the total number of items in that variable. In addition, the one-sample t test was used to determine whether the variables were significantly different from zero. Furthermore, a One-Way Analysis of Variance (ANOVA) was used to see whether there was any significant difference within different variables. A bivariate correlation was also used to determine the strength of correlations among the variables.

Results

This synthesis reviewed 23 studies using GFCF diets, for a total number of 462 individuals with ASD. These studies were selected from 15 peer-reviewed journals published from 1977 to 2010. Results of both descriptive and statistical analyses were reported across the participants' characteristics, characteristics of the dependent variables (DV), features of the independent variables (IV), and features of the study.

Articles Selected for Inclusion

Among the studies selected for analysis, thirteen studies (56.5%) were published between 2000 and 2010. Furthermore, the 23 studies selected in this research synthesis were published in 15 journals. Four studies were published in the journal, *Nutritional Neuroscience* (17.4%), and three in the journal, *Focus on Autism and Other Developmental Disorders* (13%). In addition, each of the following three journals included two studies (8.7%): *Brain Dysfunction, Journal of Autism and Childhood Schizophrenia,* and *Journal of Autism and Developmental Disorders.* Table 1 presents the frequency and percentage of the selected studies published by journal.

Descriptive Analyses by Participants' Characteristics

There were 462 individuals with ASD who participated in these 23 studies as the target in-

TABLE 1

Frequency of Reviewed Articles by the Journal

Number	*Journal*	*Frequency*	*Percent (%)*
1	Nutritional Neuroscience	4	17.4
2	Focus on Autism and Other Developmental Disabilities	3	13.0
3	Brain Dysfunction	2	8.7
4	Journal of Autism and Childhood Schizophrenia	2	8.7
5	Journal of Autism and Developmental Disorders	2	8.7
6	Autism	1	4.35
7	Behavioral Interventions	1	4.35
8	Journal of Alternative and Complementary Medicine	1	4.35
9	Journal of Applied Nutrition	1	4.35
10	Journal of Child Neurology	1	4.35
11	Journal of Endocrine Genetics	1	4.35
12	Journal of Human Nutrition Dietetics	1	4.35
13	Journal of Medical Speech-Language Pathology	1	4.35
14	Panminerva Medica	1	4.35
15	Scandinavian Journal of Educational Research	1	4.35
Sum		**23**	**100**

dividuals, with ages ranging from two to twenty-one years. Eighteen studies included 294 males (63.6%) as the target individuals, compared to fourteen studies with 73 females (15.8%). Another six studies with 95 target individuals did not specify the gender of the participants (20.6%). Ten studies included 113 target individuals diagnosed with autism (24.5%), and three studies included 76 individuals with ASD (16.5%), while no study included target children diagnosed with Rett Syndrome or Childhood Disintegrative Disorder (CDD). In four studies, participants (49.1%; $n = 227$) were referred by professionals or parents, and ten studies included individuals (31.4%; $n = 145$) representing samples of convenience, while another nine studies (with 90 participants; 19.5%) did not report the recruitment procedure. Table 2 presents the frequency of studies, number and percentage of the participants analyzed by the participants' characteristics, such as gender, diagnosis, and selection process.

Descriptive Analyses by Features of the Dependent Variables

The selected 23 studies used GFCF diets to increase communication skills (e.g., eye contact, vocalization, echolalia, nonverbal communication), social interaction, motor abilities, or cognitive abilities, or to decrease challenging behaviors (e.g., tantrums, pica, self-injury, physical aggression, property destruction, stereotypy, play, food consumption, food rejection, gagging, escape). Some of the studies also reported biomedical results to evaluate the efficacy of the intervention (e.g., Urinary Peptide Levels (UPL), gastrointestinal symptoms, Ig-antibody levels).

The behavioral variables were measured in various ways: direct observations; anecdotal parent and/or professional reports; parental and/or teacher interviews; behavior questionnaires; parent/teacher behavior ratings [DIPAB: the Diagnose of Psykotisk Adfored hos Born (Diagnosis of Psychotic Behaviour in Children)]; and parental satisfaction scale. The following standardized tests were used to measure behavioral, communicative, linguistic, cognitive, motor skills: Childhood Autism Rating Scale (CARS), Autistic Behaviour Summarized Evaluation Scale (BSE), Autism Diagnostic Observation Schedule (ADOS), Gilliam Autism Rating Scale (GARS), Vineland Adaptive Behavior Scales (VABS), Attention-Deficit Hyperactivity Disorder–IV scale (ADHD-IV), Ecological Communication Orientation (ECO) Language Sampling Summary, The Illinois Test of Psycholinguistic Abilities (ITPA), Leiter

TABLE 2

Data Related to the Participants' Characteristics

Variables	*Frequency of Studies*	*Number of Participants*	*Percent (%)***
Target Individual's Gender			
Male	18	294	63.6
Female	14	73	15.8
Not Specified	6	95	20.6
SUM	**38***	**462**	**100**
Target Individual's Diagnosis			
ASD	3	76	16.45
Autism	11	114	24.7
ASD, Autism, Asperger's Syndrome or Pervasive Developmental Disorder–Not Otherwise Specified (PDD-NOS)	4	236	51.1
Combination	5	36	7.75
SUM	**23**	**462**	**100**
Participant's Selection Process			
Referred	4	227	49.1
Convenience Sample	10	145	31.4
Not Specified	9	90	19.5
SUM	**23**	**462**	**100**

Note: * The total number of the studies according to the target individual's gender exceeds 23 due to the fact that one study may include both males and females.

** The percentage is based on the number of participants instead of the frequency of the studies.

Nonverbal Intelligence Test, Leiter International Performance Scale, the Reynells' Sprak Test, Movement Assessment Battery for Children, Kaufmann Assessment Battery for Children, and C-Raven Progressive Matrices, Tajford Observation Scheme. Biomedical tests were also used to assess urine, skin, and blood change: Urine analysis (HPLC-Gradient elution high performance liquid chromatography), skin tests (Prick tests), and blood tests of antibodies. Six studies used direct observations (26.1%). Four studies used anecdotal parent and/or professional report (17.4%), and two used standardized tests (8.7%). Almost half the studies ($n = 11$) combined more than one way to measure the change of the dependent variables (47.8%).

Eight studies described the dependent variables with an operational definition, using observable and measurable terms (34.8%); 11 studies did not use operational definitions (47.8%). Another four studies reported both medical and behavior variables, with medical variables included in operational definitions, but without behavioral variables included in operational definitions (17.4%). Ten studies used quantifiable measurement to describe the dependent variables with numbers (43.5%); nine studies did not use a quantifiable index (39.1%), and another four studies mixed quantifiable results with anecdotal reports (17.4%). Ten studies described the measurement procedure in valid, precise and replicable terms so that other researchers may replicate the procedure in similar studies (43.5%); nine studies did not use valid, precise, and replicable procedures (39.1%), and another four studies mixed the two descriptions (17.4%).

More than half the studies ($n = 16$) did not report maintenance or generalization data and procedures across participants (69.6%). Among the seven studies that reported maintenance, the follow-up period ranged from six months to eight years, and only one study reported generalization across participants (4.3%). The majority of the studies ($n = 17$) did not report interrater reliability (73.9%). Among the six studies that reported interrater reliability, all percentages were higher than

TABLE 3

Data Related to the Features of the Dependent Variables

Variables	*Frequency*	*Percent (%)*
Means of Measurement		
Anecdotal Report	4	17.4
Direct Observation	6	26.1
Standardized Tests	2	8.7
Combination	11	47.8
Operational Definition		
Yes	8	34.8
No	11	47.8
Mixed	4*	17.4
Quantifiable Measurement		
Yes	10	43.5
No	9	39.1
Mixed	4*	17.4
Valid and Precise Procedure		
Yes	10	43.5
No	9	39.1
Mixed	4*	17.4
Maintenance and/or Generalization		
Yes	7	30.4
No	16	69.6
Interrater Reliability		
Yes: Higher than .9	5	21.7
Yes: Higher than .8	1	4.4
Yes: Sum	**6**	**26.1**
No	17	73.9
Social Validity		
Yes: Mentioned and Reported Data	0	0
No: Mentioned but no Data Reported	2	8.7
No: Not Mentioned	21	91.3
No: Sum	**23**	**100**

Note: * A mixed study included multiple variables, including medical variables and behavioral variables.

0.80, which were acceptable. Five of them reported interrater reliability percentages higher than 0.9 (21.7%). Furthermore, most of the studies ($n = 21$) did not mention social significance of the study (91.3%). Only two mentioned social validity, yet did not report the data (8.7%). Table 3 presents the frequency and percentage of the studies analyzed by the features of the dependent variables discussed above.

Descriptive Analyses by Features of the Independent Variables

Eleven studies used GFCF diets (47.8%), and one used a gluten-free diet (4.4%). Another eleven studies used multiple interventions in addition to GFCF diets (47.8%), including vitamin therapy-multivitamin supplements, elimination of certain foods, alternative medical therapy (CAM), environmental control and avoidance of triggers (mites, moisture, mold, smoke, pesticides, toxic cosmetics/cleaners), gastrointestinal support, antigen injection therapy, behavior intervention, and special education services (e.g., speech language pathology, occupational therapy, physical therapy).

Nine studies were described with replicable precision (39.1%), while twelve (52.2%) did not meet this standard. Another two studies included multiple interventions, and some

were described with replicable precision while some were not (8.7%). Similarly, nine studies were implemented with systematic manipulation and under the researchers' control (39.1%), while twelve (52.2%) did not meet this standard. Another two studies included multiple interventions that were implemented with mixed methods, and some of the IVs within these studies were systematically manipulated while some were not (8.7%). The majority of the studies did not mention treatment integrity/fidelity (n = 22; 95.7%). Only one study reported treatment integrity, with the data ranging from 0.95 to 1.0 (4.3%).

Seven studies used group comparison designs (30.4%), and two studies used single subject designs (one multiple baseline across behaviors and one BABA design; 8.7%). Two were AB case study designs (8.7%), and five were AB across participants designs (21.8%). Another seven studies did not identify the specific research design (30.4%).

The length of interventions ranged from 8 days to 48 months (M = 12.6 months). (The total number of the studies according to the length of the intervention exceeds 23 due to the fact that one study may include more than one case, each lasting various periods of time.) Half the studies were completed within three months (n = 13; 50%), and another seven studies lasted one year or less (26.9%). Only five studies lasted over one year (19.2%), and one study did not specify how long the intervention lasted (3.9%). The majority of the studies reported positive results (64.3%). Four reported negative results (14.3%), and six reported no significant changes (21.4%). (Similarly, the total number of studies according to the results of the intervention exceeds 23 due to the fact that one study may include more than one case, each with various results.) Table 4 presents the frequency and percentage of the selected studies analyzed by different features of the independent variables.

Statistical Analyses

Using one-sample t-tests, any significant difference between the variable and 0 was found as follows: (a) interrater reliability (t = 2.712; $p < .05$), (b) duration of the study (t = 7.713; $p = .000$), and (c) the result of the study (t = 5.163; $p = .000$). No significant difference between treatment fidelity and 0 (t = 1.000, $p > .05$) was found.

A one-way ANOVA was used to investigate whether there was a significant difference within each variable by the result of the study. There were four variables that indicated statistical significant difference by the results of the studies: (a) selection process (F = 4.890; $p < .05$), (b) interrater reliability (F = 4.095; $p < .05$), (c) social validity (F = 3.304; $p < .05$), and (d) the length of the intervention (F = 6.064; $p < .01$).

Significant, positive correlations were found between quantifiable measurement and valid and precise procedure (r = 1.000; $p = .000$), as well as between replicable procedure and systematic manipulation (r = 1.000; $p = .000$). Operational definition of the dependent variable had a significant, positive correlation with quantifiable measurement (r = .965; $p = .000$) and valid and precise procedure (r = .965; $p = .000$). The measurement means of the dependent variables were significantly correlated with operational definitions (r = .462; $p < .05$), quantifiable measurement (r = .521; $p < .05$), and valid and precise procedure (r = .521; $p < .05$). Replicable procedure and systematic manipulation of the intervention had significant correlations with the operational definition (r = .559; $p < .01$), quantifiable measurement (r = .610; $p < .01$), valid and precise procedure (r = .610; $p < .01$), interrater reliability (r = .475; $p < .05$), and research design (r = .439; $p < .05$). In addition, a significant correlation existed between social validity and treatment fidelity (r = .691; $p = .000$) and between the length of the study and the measurement means of the dependent variable (r = .518; $p < .05$). There was, however, no bivariate correlation between the results of the study and different variables ($p > .05$). Table 5 presents the correlation r and ANOVA F scores with p values of the dependent and independent variables.

Discussion

The authors conducted a research synthesis across 23 studies using GFCF diet interventions from 15 peer-reviewed journals published between 1977 and 2010. There was no study published before 1977 that met the se-

TABLE 4

Data Related to the Features of the Interventions

Variables	*Frequency*	*Percent (%)*
Intervention		
GFCF	11	47.8
GF	1	4.4
Multiple Interventions	11	47.8
Replicable Description		
Yes	12	52.2
No	9	39.1
Mixed	2*	8.7
Systematic Manipulation		
Yes	12	52.2
No	9	39.1
Mixed	2*	8.7
Treatment Fidelity		
Yes	1	4.3
No	22	95.7
Research Design		
Group Comparison Design	7	30.4
Single Subject Design	2	8.7
AB Design	7	30.4
Unidentified	7	30.4
Length of the Intervention		
Within 1 month (including 1 month)	7	26.9
1 ~ 3 months (including 3 months)	6	23.1
3 ~ 6 months (including 6 months)	2	7.7
6 ~ 12 months (including 12 months)	5	19.2
Over 12 months	5	19.2
Unidentified	1	3.9
SUM	**26****	**100**
Results of the Intervention		
Positive	18	64.3
Negative	4	14.3
No Significant Changes	6	21.4
SUM	**28*****	**100**

Note: * A mixed study included multiple variables, including medical variables and behavioral variables.

** The total number of the studies according to the length of the intervention exceeds 23 due to the fact that one study may include more than one case which lasted various period of time.

*** Similarly, the total number of the studies according to the results of the intervention exceeds 23 due to the fact that one study may include more than one case which had various results.

lection criteria. Among the 462 individuals with ASD, no individuals were reported with Rett Syndrome or CDD.

Almost one -fifth of the studies did not mention the selection procedure (n = 9; 19.5%). Almost half the studies did not describe the dependent variables with an adequate operational definition (n = 11, 47.8%), and two-fifths were without a measurable definition (n = 9, 39.1%), while two-fifths were without a valid and precise procedure description (n = 9, 39.1%). Lack of detailed and precise information will make it very difficult for other researchers to replicate these studies.

The operational definition standard had strong correlations with quantifiable measurement (p = .000) and valid and precise procedure (p = .000). Quantifiable measurement had a strong bivariate correlation with valid and precise procedure (p = .000). In addi-

TABLE 5

Correlations and ANOVA *F* Scores and *p* Values on Dependent Variables and Independent Variables

	Correlation r (p)				
	Operational Definition	*Quantifiable Measurement*	*Valid and Precise Procedure*	*Interrater Reliability*	*Research Design*
Operational Definition		.965** (.000)	.965** (.000)		
Means of Measurement	.462* (.026)	.521* (.011)	.521* (.011)		
Replicable Procedure	.559** (.006)	.610** (.002)	.610** (.002)	.475* (.022)	.439* (.036)
Systematic Manipulation	.559** (.006)	.610** (.002)	.610** (.002)	.475* (.022)	.439* (.036)

	One-Way ANOVA by the Results of the Study F *and p*	
	F	*p*
Selection Process	4.890▪	.011
Interrater Reliability	4.095▪	.021
Social Validity	3.304▪	.043
Length of the Intervention	6.604▪▪	.004

Note: * indicates that the correlation between the two variables is significant at $p < .05$ level;
** indicates that the correlation between the two variables is significant at $p < .01$ level.
▪ indicates that there is significant difference within each variable at $p < .05$ level;
▪▪ indicates that there is significant difference within each variable at $p < .01$ level;

tion, positive bivariate correlations were found between the means of the measurement and the operational definition, quantifiable measurement, valid and precise procedure, and the length of the intervention ($p < .05$). Operational definition, quantifiable measurement and valid and precise procedure are contributors to a study with strong design.

Seventy percent of the studies did not report maintenance, which provided no evidence of the efficacy of the intervention in the long term ($n = 16$). Among 23 studies, only one reported generalization across participants (4.3%). The field needs more studies that include maintenance and generalization procedures and data, yet the body of the existing GFCF diet intervention studies failed to provide it.

Three-fourths of the studies did not report interrater reliability ($n = 17$; 73.9%), and 91.3% of the studies did not mention social validity ($n = 21$). The two studies that did mention social validity did not report these data ($n = 2$; 8.7%).

Result of the one-way ANOVA indicated that there was a significant difference within the variable of interrater reliability by the result of the study ($p < .05$). Similarly, a significant difference was also found within the variable of social validity by the result of the study ($p < .05$). The ultimate purpose of the intervention was to improve the quality of life of the individuals with ASD through the improvement of skills, abilities, appropriate behaviors and the decrease of the inappropriate behaviors. Yet most of the studies failed to evaluate the social significance of the changes the interventions brought into families' and participants' lives.

Almost half the studies included multiple interventions in addition to GFCF diets ($n = 11$; 47.8%), which led to the uncertainty regarding the contribution of the GFCF diet intervention to the results of the studies. Half the studies did not describe the interventions in replicable or systematic terms ($n = 12$; 52.2%), which makes it very difficult, if not impossible, for other researchers to replicate these studies. The majority of the studies did not report treatment fidelity ($n = 22$; 95.7%) and lack of treatment fidelity analysis means lack of evidence that the researchers conducted the intervention and measured the de-

pendent variables as the study purported they were conducted and measured.

Replicable precision had a strong bivariate correlation with systematic manipulation ($p = .000$). In addition, both replicable precision and systematic manipulation had positive correlations with operational definition, quantifiable measurement, valid and precise procedure of the dependent variables, interrater reliability, and research design. Similar to the correlations among the indicators of the dependent variables, these are all contributors to a study with strong design.

Almost one-third of the studies failed to identify the specific research design ($n = 7$; 30.4%). Half the studies conducted the intervention within three months ($n = 13$, 50%), while only five studies lasted longer than one year (19.2%). Result of the one-way ANOVA indicated that there was a significant difference within the variable of length of the intervention by the result of the study ($p < .01$). Thus, the longer the study lasted, the better the result and the more questionable the internal validity of the study, e.g., due to threats such as history and maturation.

Four studies reported negative results (14.3%), and one-fifth reported no significant changes ($n = 6$, 21.4%), while more than half the studies reported positive results ($n = 18$; 64.3%). No statistically significant correlations were found between the results of the studies and the other variables ($p > .05$). However, since only studies published in peer-reviewed journals were included, this synthesis was biased in favor of published journal papers. Due to a potential bias imposed by publication procedures, studies with negative effects or with no significant changes are less likely to be submitted and published compared to the ones with positive results (Horner, Carr, Strain, Todd, & Reed, 2002). This may affect the accuracy of the synthesis result.

Implications

Even though this research synthesis has its limitations due to the inclusion criteria utilized (e.g. peer-reviewed journal papers, most of which reported positive results), it still brings up several suggestions to be considered for future studies using GFCF diets for individuals with ASD. More research needs to be conducted with individuals with Rett Syndrome and CDD on the efficacy of GFCF diet interventions, though the prevalence of these two categories is admittedly low.

Almost half the studies combined GFCF diets with other intervention components, which made it difficult to accurately analyze the effectiveness of the GFCF diet interventions. More studies need to implement only GFCF diets for individuals with ASD so that evidence can be provided to indicate whether GFCF diet interventions are truly effective. In addition, more studies conducted across longer periods of time, especially over one year, need to be carried out, but with more stringent controls for threats to internal validity. Due to the complexity of the human digestive and other internal systems, it is difficult to pinpoint if GFCF diet interventions are the only factor that contributes to a positive result in these studies.

More studies with research designs that better reduce error, especially rigorous single subject designs, need to be implemented. Results of the statistical analyses indicate that operational definition and quantifiable measurement of the dependent variables, valid and precise procedure, replicable procedure and systematic manipulation of the intervention are all positively correlated with each other. They are all factors that contribute to the sound design of a study, yet over half the studies did not adequately report these factors. Similarly, most studies failed to report interrater reliability, social validity, and treatment fidelity, though they are essential to insure the reliability and validity of a study. More studies need to also consider maintenance and generalization, since they are helpful to clarify the sustainability of an intervention. All these reliable, valid, and precise descriptions will facilitate replication for future researchers and contribute to the literacy of evidence-based practice, thus generating a greater impact for individuals with ASD, their families, and the professionals working with them.

Results of this research synthesis agree with the existing literature (National Autism Center, 2009) on the efficacy of GFCF diet interventions: There is little scientific evidence to draw a firm conclusion that the GFCF diets intervention is effective for individuals with

ASD. With the consideration of the amount of expense and time that GFCF diet interventions tend to involve, researchers in the future need to implement better controlled studies with more objective assessment tools to evaluate more accurately the efficacy of the intervention.

References

References marked with an asterisk (*) indicate studies included in the research synthesis.

*Adams, L., & Conn, S. (1997). Nutrition and its relationship to autism. *Focus on Autism and Other Developmental Disabilities, 12,* 53–64.

*Bird, B., Russo, D., & Cataldo, M. (1977). Considerations in the analysis and treatment of dietary effects on behavior: A case study. *Journal of Autism and Childhood Schizophrenia, 7,* 373–382.

*Bowers, L. (2002). An audit of referrals of children with autistic spectrum disorder to the dietetic service. *Journal of Human Nutrition Dietetics, 15,* 141–144.

*Cade, R., Privette, M., Fregly, M., Rowland, N., Sun, Z., Zele, V., et al. (2000). Autism and Schizophrenia: Intestinal disorders. *Nutritional Neuroscience, 3,* 57–72.

Centers for Disease Control and Prevention. (2007). *Prevalence of ASDs.* Retrieved from http://www.cdc.gov/ncbddd/autism/fag_prevalence.htm.

Centers for Disease Control and Prevention. (2010). *Prevalence of ASDs.* Retrieved from http://www.cdc.gov/ncbddd/autism/fag_prevalence.htm.

Christison, G. W., and Ivany, K. (2006). Elimination diets in autism spectrum disorders: Any wheat amidst the chaff? *Developmental and Behavioral Pediatrics, 27,* 162–171.

Council for Exceptional Children–Professional Standards & Practice Committee. (2006). *CEC evidence-based professional practices proposal.* Retrieved from http://www.cec.sped.org/Content/NavigationMenu/ProfessionalDevelopement/ProfessionalStandards/EVP_REVISED_03_2006.pdf.

*Elder, J. H., Shankar, M., Shuster, J., Theriaque, D., Burns, S., & Sherrill, L. (2006). The gluten-free, casein-free diet in autism: Results of a preliminary double-blind clinical trial. *Journal of Autism and Developmental Disorders, 36,* 413–420.

*Genuis, S. J., & Bouchard, T. P. (2010). Celiac disease presenting as autism. *Journal of Child Neurology, 25,* 114–119.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C, & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71,* 149–164.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single subject research to identify evidence-based practice in special education. *Exceptional Children, 71,* 165–179.

Horner, R. H., Carr, E. G., Strain, P. S., Todd, A. W., & Reed, H. K. (2002). Problem behavior interventions for young children with autism: A research synthesis. *Journal of Autism and Developmental Disorders, 32,* 423–446.

Individuals with Disabilities Education Improvement Act of 2004. (IDEiA, 2004). P.L. 108-446, 20 U.S.C. §1400 et seq.

*Irvin, D. (2006). Using analog assessment procedures for determining the effects of a gluten-free and casein-free diet on rate of problem behaviors for an adolescent with autism. *Behavioral Interventions, 21,* 281–286.

*Knivsberg, A., Wiig, K., Lind, G., Nodland, M., & Reichelt, K. (1990). Dietary intervention in autistic syndromes. *Brain Dysfunction, 3,* 315–327.

*Knivsberg, A., Reichelt, K., Nodland, M., & Hoien, T. (1995). Autistic syndromes and diet: A follow-up study. *Scandinavian Journal of Educational Research, 39,* 223–236.

*Knivsberg, A., Reichelt, K., & Nodland, M. (1999). Dietary intervention for a seven year old girl with autistic behavior. *Nutritional Neuroscience, 2,* 435–439.

*Knivsberg, A., Reichelt, K., Hoien, T., & Nodland, M. (2002). A randomized, controlled study of dietary intervention in autistic syndromes. *Nutritional Neuroscience, 5,* 251–261.

*Knivsberg, A, Reichelt, K., Hoien, T., & Nodland, M. (2003). Effect of a dietary intervention on autistic behavior. *Focus on Autism and Other Developmental Disorders, 18,* 247–256.

*Lucarelli, S., Frediani, T., Zingoni, M., Ferruzzi, F., Giardini, O., Quintieri, F., et al. (1995). Food allergy and infantile autism. *Panminerva Medica, 37,* 137–141.

Mayton, M. R., Wheeler, J. J., Menendez, A. L., & Zhang, J. (2010). An analysis of evidence-based practices in the education and treatment of learners with autism spectrum disorders. *Education and Training in Autism and Developmental Disabilities, 45,* 539–551.

Munasinghe, S. A., Oliff, C., Finn, J., & Wray, J. A. (2010). Digestive enzyme supplementation for autism spectrum disorders: A double-blind randomized controlled trial. *Journal of Autism and Developmental Disorders, 40,* 1131–1138.

National Autism Center. (2009). *Findings and conclusions: Addressing the need for evidence-based guidelines for autism spectrum disorders.* Retrieved from http//www.nationalautismcenter.org.

No Child Left Behind Act (Pub. L. No. 107-110)

(2001). Washington, DC: U.S. Government Printing Office.

*O'Banion, D., Armstrong, B., Cummings, R., & Stange, J. (1978). Disruptive behavior: A dietary approach. *Journal of Autism and Childhood Schizophrenia, 8,* 325–337.

*Patel, K., & Curtis, L. (2007). A comprehensive approach to treating autism and attention-deficit hyperactivity disorder: A pre-pilot study. *Journal of Alternative and Complementary Medicine, 13,* 1091–1097.

*Reichelt, K., Ekrem, J., & Scott, H. (1990). Gluten, milk proteins and autism: Dietary intervention effects on behavior and peptide secretion. *Journal of Applied Nutrition, 42,* 1–10.

*Reichelt, K. L., Knivsberg, A., Lind, G., & Nodland, M. (1991). Probable etiology and possible treatment of childhood autism. *Brain Dysfunction, 4,* 308–319.

*Seung, H., Rogalski, Y., Shankar, M., & Elder, J. (2007). The gluten- and casein-free diet and autism: Communication outcomes from a preliminary double-blind clinical trial. *Journal of Medical Speech-Language Pathology, 15,* 337–345.

*Shabo, Y., & Yagil, R. (2005). Etiology of autism and camel milk as therapy. *Journal of Endocrine Genetics, 4,* 67–70.

Wheeler, J. J. (2007, November). *Evidence-based practice in the treatment of autism spectrum disorders: Implications for children, families and professionals.* Keynote Speaker at the West Virginia State Conference on Autism, West Virginia University, Morgantown, WV.

*Whiteley, P., Haracopos, D., Knivsberg, A., Reichelt, K. L., Parlar, S., & Jacobsen, J., et al. (2010). The ScanBrit randomized, controlled, single-blind study of a gluten- and casein-free dietary intervention for children with autism spectrum disorders. *Nutritional Neuroscience, 13,* 87–100.

*Whiteley, P., Rodgers, J., Savery, D., & Shattock, P. (1999). A gluten-free diet as an intervention for autism and associated spectrum disorders: Preliminary findings. *Autism, 3,* 45–66.

Whiteley, P., & Shattock, P. (2002). Biochemical aspects in autism spectrum disorders: Updating the opioid-excess theory and presenting new opportunities for biochemical intervention. *Expert Opinion on Therapeutic Targets, 6,* 175–183.

*Wong, H. H. L., & Smith, R. G. (2006). Patterns of complementary and alternative medical therapy use in children diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36,* 901–909.

*Wood, B. K., Wolery, M., & Kaiser, A. P. (2009). Treatment of food selectivity in a young child with autism. *Focus on Autism and Other Developmental Disorders, 24,* 169–177.

Zhang, J., & Wheeler, J. J. (2011). A meta-analysis of peer-mediated interventions for young children with autism spectrum disorders. *Education and Training in Developmental Disabilities, 46,* 62–77.

Received: 5 January 2012
Initial Acceptance: 1 March 2012
Final Acceptance: 10 May 2012