Healthy nutrition in Germany: A survey analysis of social causes, obesity and socioeconomic status

Sebastian Mader[1*], Malte Rubach[2], Wolfram Schaecke[2], Christine Röger[3], Ina Feldhoffer[3] and Eva-Magdalena Thalmeier[3]

[1] Institute of Sociology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland, Email: Sebastian.Mader@soz.unibe.ch

[2] Bavarian State Ministry for Nutrition, Agriculture, and Forestry, Ludwigstrasse 2, 80539 Munich, Germany

[3] Competence Centre of Nutrition (KErn) at the Bavarian State Research Center for Agriculture, Am Gereuth 4, 85354 Freising, Germany

* Corresponding author

21 November 2019

Acknowledgements: Sebastian Mader thanks Prof. Dr. Axel Franzen (Institute of Sociology, University of Bern), Dr. Gert Mensink (Robert Koch-Institute Berlin) and two anonymous referees for helpful comments and suggestions, and the Max Rubner-Institute and Dr. Gert Mensink to make the datasets available and accessible for analysis.

Financial Support: This study was funded by the Bavarian State Ministry for Nutrition, Agriculture, and Forestry (StMELF) within the project 'Nutrition in Bavaria and its Costs' ('Ernährungsverhalten in Bayern und seine Folgekosten', project number A/15/32) of the Competence Centre of Nutrition (KErn) at the Bavarian State Research Center for Agriculture. NVS II and NEMONIT were funded by the German Federal Ministry of Food, Agriculture and Consumer Protection.

Conflict of interest: none.

Authorship: S.M. developed the concept of the paper, analyzed the data, and wrote a first draft of the manuscript. M.R., W.S., and C.R. also co-developed the concept of the paper and revised the manuscript. I.F. and E.-M.T. also revised the manuscript.

Ethical Standards Disclosure: not applicable.

Keywords: nutritional behavior research; nutrition-related health risks; obesity; mixed effects panel regression;

This is a post-peer-review, pre-copyedit version of an article published in Public Health Nutrition on 27 April 2020: DOI: 10.1017/S1368980019004877.

Healthy nutrition in Germany: A survey analysis of social causes, obesity and socioeconomic status

Abstract

Objective: The obesity pandemic is an increasing burden for society. Information on key drivers of the nutrition cycle of (a) social causation, (b) biological causation, and (c) health selection is vital for effective policies targeted at the reduction of obesity prevalence. However, empirical causal knowledge on (a) the social predictors of diet quality, (b) its impact on corpulence, and (c) the socio-economic consequences of obesity is sparse. We overcome the limitations of previous research and acquire comprehensive causal insight into this cycle.

Design: Therefore, we analyze two German socio-epidemiological panel surveys exploiting their longitudinal panel structure utilizing hybrid panel regression models.

Setting: general population of Germany.

Participants: German Health Interview and Examination Survey for Children and Adolescents (KiGGS, n=17’640, age: 0-24) and the German National Nutrition Monitoring (NEMONIT, n=2’610, age: 15-82).

Results: The results indicate that (a) interestingly only gender, education, and age explain healthy diets. (b) Increases in a newly developed Optimized Healthy Eating Index (O-HEI-NVSII), and in nuts intake reduce body mass index, while growing overall energy intake, lemonade, beer, and meat (products) intake drive corpulence. (c) In turn, developing obesity decreases socio-economic status.

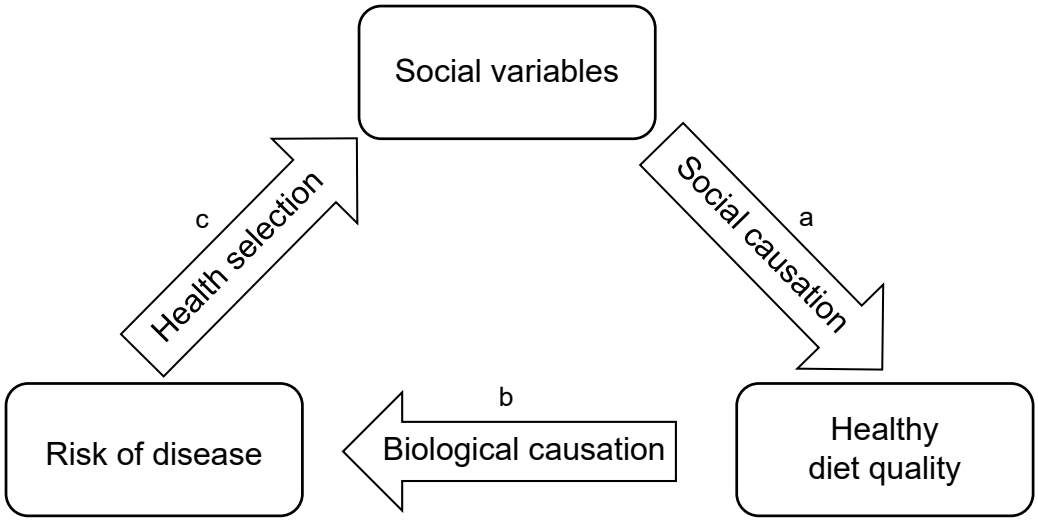
Conclusions: These results suggest that policies targeted at the reduction of obesity prevalence may be well advised to focus on boys and men, people with low education, the promotion of a healthy diet and nuts intake, and the limitation of lemonade, beer, and meat (products) intake. Therefore, future research may focus on the replication of our findings utilizing longer panels and experimental approaches.

Not only in developed countries, the obesity pandemic and its related secondary diseases are an increasing burden for social care systems, economic productivity, and individual quality of life[1]. People suffering from overweight and obesity have a higher risk of dying prematurely[2], developing cardiometabolic multimorbidity[3], Alzheimer's disease[4], and pancreatic and breast cancer[5,6]. With its unrelieved growth of about 1% per year[7], the obesity pandemic puts a serious and increasing burden of direct and indirect costs on social security systems as well as tangible and intangible costs on individuals[1,8]. In contrast, a balanced diet is expected to reduce obesity-induced morbidities and all-cause mortality[9].

Various theoretical approaches have been used to explain diet quality, and its health outcomes[10,11]. The two most prominent hypotheses derived from them are the 'social causation hypothesis' and the 'health selection hypothesis'. While the former theorizes that manifold socio-economic, -demographic, and -cultural factors drive health outcomes of diets like corpulence, the health selection hypothesis inter alia assumes that nutritional health outcomes predict social outcomes. So far, evidence for the relative empirical validity of these seemingly opposing hypotheses is inconclusive and fragmented[12]. Moreover, the social causation hypothesis often implicitly assumes a direct deterministic link between diet quality and corpulence. For that reason, most studies testing the social causation hypothesis directly regress nutritional health outcomes on social factors, e.g. body mass index (BMI) on socio-economic status (SES). However, for the identification of the causal nutritional socio-epidemiological mechanisms this approach over-simplifies, as it disregards the biological causation path regressing health outcomes on diets. Only the inclusion of this path guarantees comprehensive insight in the causal relationships in a so-called 'nutrition cycle' (see Figure 1), where social factors are hypothesized to predict diet quality (social causation (a)), nutrition impacts corpulence (biological causation (b)), and corpulence affects social outcomes (health selection (c)). This causal spiral in time can be viewed both *intra*-generational explaining intra-personal life-courses and *inter*-generational referring to the (epi)genetic and social inheritance of health chances.

Social determinants of diet quality (a) are hypothesized to be manifold[10,13,14]. They range from individual and contextual socio-economic factors (e.g. SES, employment status, healthy food availability) to socio-demographic factors (e.g. sex, age, migration background) to socio-cultural factors (e.g. gender, migration background and regional

Figure 1: The nutrition cycle.



diet imprints of origin) and other factors (e.g. physical activity and psychological problems)[13]. In this vein, social factors affecting nutrition outcomes could also be subsumed to the categories affordability, availability, and accessibility of foods. This so-called 'triple-A' model[14] is a synthesis of the principles of behavioral economics and human-ecological setting approaches[15] to model the dimensions of individual dietary decisions: Affordability comprises both direct costs and opportunity costs of foods. Availability refers to the availability of foods, as well as to the surrounding opportunity structure, and thus to the contextual features of a social space. Finally, accessibility encompasses internalized cultural knowledge and the associated scripts. More specifically, accessibility refers to characteristics like educational background or attainment, as well as to normative attitudes and associated characteristics as regards the socio-economic and socio-cultural attributes of foods[16].

Generally, information on key drivers of the 'nutrition cycle' is vital for effective explicit and implicit behavioral political interventions targeted at the reduction of obesity prevalence and health promotion[17,18]. Yet, causal empirical evidence on the 'nutrition cycle' is sparse. So far, there are only six studies that analyze population-based socio-epidemiological panel surveys with measurements of all relevant variables at all points in time for either (a)[19,20], or (b)[21,22,23], or (c)[24] and applying panel regression models (see Table 1).

Table 1: Summary of the state of research for the nutrition cycle applying causal inference.

	Outcome	Dim.	Covariate	Eff.	FE	NPS	CAR	Country	n	Period	Study	LoC
Social Causation	diet quality	Afford.	income	→	yes	yes	no	U.S.	8136	2008-2012	Zagorsky and Smith (2017)[20]	-
			wealth	u								
		Avail.	nr. of food outlets in neighborhood	↗	yes	no	no	U.S.	1053	1985-2011	Rummo et al. (2017)[19]	o
			move into a city	↘	yes	yes	no	U.S.	8136	2008-2012	Zagorsky and Smith (2017)[20]	-
		Access.	education	↗	no	various cross-sectional studies (see e.g. Fekete and Weyers 2016[13] for an overview).						-
Biological Causation	BMI		diet quality	↘	yes	no	no	BE	570	2002-2014	Mertens et al. (2015)[21]	-
				↘	no	no	no	U.S.	146071	1986-2006	Fung et al. (2015)[22]	-
	WC			↘	yes	yes	no	U.S.	67175	1993-1998	Cespedes Feliciano et al. (2016)[23]	-
Health Selection	SES		BMI	↘	yes	yes	no	U.S.	3386	1992-2002	Michaud and van Soest (2008)[24]	-

Note: Dim. = Dimension, Eff. = Effect, FE = fixed effects panel regression, NPS = national panel survey, CAR = complete age range, n = number of subjects, LoC = level of confidence, Afford. = Affordability, Avail. = Availability, Access. = Accessibility, → = no effect, ↗ = positive effect, ↘ = negative effect, u = u-shaped effect, - = low, o = middle, BMI = body mass index, WC = waist circumference, SES = socio-economic status, U.S. =United States of America, BE = Belgium;

All six presented studies advantageously use panel data applying multiple panel regression models. Nonetheless and besides individual weaknesses, these six studies have some limitations in common: First, five of the studies are based on data from the U.S.[19,20,22,23,24]; only one study stems from another country (Belgium)[21]. Hence, the state of research on the nutrition cycle is geographically restricted to two countries. Second, none of the six studies includes time-dummies as confounders that would to account for overall time-trends (i.e. overall time-varying unobserved heterogeneity). Third, none of the six studies covers the full age range. Fourth and finally, all of the six studies are limited to the analysis of only one of the analytical paths of the nutrition cycle ((a) or (b) or (c)).

Data and methods

Hence, this study aims at overcoming these limitations and acquires comprehensive insight into the whole nutrition cycle by analyzing two German socio-epidemiological survey panel datasets utilizing hybrid generalized linear mixed panel regression models: two waves of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) aged 0 to 24 years (n=17'640) between 2003 and 2012, and five waves of the German National Nutrition Monitoring (NEMONIT)[25] panel study for 15- to 82-year-olds (n=2'610) surveyed between 2005-2013. Both the KiGGS and NEMONIT datasets consist of random samples of the German population and together cover an age range from 0 to 82 years as well as an observation period of about a decade between 2003 and 2013. KiGGS 0 received ethical approval from the German Federal Data Protection Office and the Ethics Committee of Charité University Medicine[26]. KiGGS 1 received ethical approval from the Hanover Medical School's ethics committee and the Federal Commissioner for Data Protection and Freedom of Information[27]. NEMONIT received ethical approval from the German Federal Data Protection Office[28]. Written informed consent was obtained in detail from the respondents in all of the three studies[26-28].The central characteristics of the KiGGS and NEMONIT panels, also with regard to the measurement of central characteristics for this study, are summarized in Table 2. The Tables S1, S2, S3, and S4 of the Supplementary Material show descriptive statistics and measurement descriptions of all variables included in the analyses.

Affordability and availability contexts

The comparatively large number of cases and sample points in the KiGGS panel allows a relatively fine-grained spatial segmentation and therefore the analysis of effects of affordability and availability contexts on diet quality: County- and state-level data on average disposable income per capita (p.c.) from the Regional Database Germany (RDG) of the statistical offices of the confederation and the federal states was merged with the KiGGS data. County-level data on the number of retail firms and firms in the food service industry p.c., and on the factory area of trade and industry p.c. was also added from the RDG. These three context indicators are used to approximate the variety of healthy food supply, as purchasing power is

Table 2: Comparison of central data characteristics

	KiGGS	NEMONIT
cases	N=29‘632, n=17‘640	N=13‘050, n=2‘610
waves	2 (KiGGS 0 + KiGGS 1)	5 (NVSII + 4*NEMONIT)
time span	2003-2012	2005-2013
age range	0-24	15-82
spatial segmentation	counties, states	regions (east, west, north, south)
food intake	food frequency questionnaire of usual intake of 5 food groups in the last 4 weeks	two 24h-recalls of acute intake of 17 food groups
anthropometry	not available for panel analysis	Body mass index (BMI)
socio-economic, -demographic, and -cultural variables	extensive	extensive

Note: KiGGS = German Health Interview and Examination Survey for Children and Adolescents, NEMONIT = German National Nutrition Monitoring, NVSII = National Nutrition Survey II; ‘n’ refers to the number of individuals, and ‘N’ to the number of observations (‘n’ times the number of years ‘T’); data sources: KiGGS Panel: Robert Koch-Institute. NEMONIT: Max Rubner-Institute (2016)[25].

an important location factor for food retailers. Moreover, the number of firms, and sales area are associated with product variety[29].

Food intake

As the concept of ‘healthy nutrition’ is widely disputed and operationalized in manifold ways, in this study we provide a diverse selection of indicators of a healthy diet. The two datasets differ in the measurement of food intake: In the KiGGS panel the usual intake of only five food groups (fruits, vegetables, juices, sweetened soft drinks, and sweets) is measured with a food frequency questionnaire retrospectively for the last four weeks[30] in both available waves. However, this is no fundamental restriction for the calculation of indicators of intake quality, for example following the concept of the Healthy Nutrition Score for Kids and Youth (HuSKY, see Kleiser et al. 2009[31] for details). HuSKY is a Healthy Eating Index (HEI) specially designed for children and adolescents in accordance with the food group specific intake recommendations of the Optimized Mixed Diet (OMD)[32,33]. In addition, an index

referring to the Healthy Food Diversity (HFD) Index[34] can be calculated from the relative recommended intake ratio of amply recommended (here: fruits, vegetables, and juices) to tolerated (i.e. foods with high energy density and low nutrient content; here: sweetened soft drinks, and sweets) food groups according to the OMD.

By contrast, in the NEMONIT panel, the acute intake of 17 food groups of the previous day is surveyed via two 24h-recall telephone interviews[35]. As the acute food intake varies by weekday and season, all models using NEMONIT diet data control for these variables. The comprehensive diet survey in NEMONIT enables the calculation of various indicators of intake quality. First, a HEI in accordance with the dietary recommendations of ten food groups of the German Nutrition Society (DGE)[36] could be computed as suggested in Gose et al. (HEI-NVSII, 2016)[28]. HEI-NVSII like other conventional Healthy Eating Indexes describes the intake compliance with dietary guidelines usually on a ratio scale from 0 = 'no compliance' to 100 = 'full compliance' based on food group specific recommendations in *absolute* grams per day (g/d). The range of the original HEI-NVSII is 0-110, as the maximum score for fruit and vegetable intake is 15 each instead of 10 for the other food groups included[28]. However, this range can be normalized to 0-100. Yet, recommendations in absolute terms depend on e.g. sex, age, body weight and level of physical activity. This introduces substantial assumptions on individual characteristics that could impose risks of producing a methodological artefact. Consequently, we suggest an Optimized HEI-NVSII (O-HEI-NVSII). This is the additive index of compliance with the dietary guidelines of ten food groups based on the sex specific *relative* shares of the whole intake (in g/d) on a scale from 0 to 100 taken from Gedrich and Karg (2001)[37]. Thus, O-HEI-NVSII supposedly has the comparative advantage that relative recommendations are only to a relatively small extent prone to individual characteristics. The single prerequisite to compute O-HEI-NVSII is information on the whole daily intake in grams, which is met by NEMONIT, but not by the KiGGS panel. Moreover, HEIs are applied to recommendations of macronutrient intake (HEI-MAC)[38] based on relative shares, and to guidelines for energy intake (HEI-EN)[39]. The HFD index is calculated following Drescher et al. (2017)[34]. Lastly, the average energy density (ED) in kcal/g is calculated. Owing to the recent discussion about the concept of ED, we compute both the ED of the caloric food intake excluding non-caloric beverages and drinking water (ED1) and the ED of the non-beverage food intake excluding all beverages (ED2)[40,41].

Anthropometry

The BMI as a measure of corpulence is only available in the NEMONIT panel. Therefore, it is possible to analyze the whole nutrition cycle with NEMONIT, while the social causation path can be analyzed for KiGGS. In NEMONIT the BMI is surveyed as self-report of body weight and height, while in the NVSII (the first wave of NEMONIT) objective measures were taken, too. However, the self-reported and the objective BMI are correlated very highly (Spearman's r_s =.98, Pearson's r_p=.99, n = 2'551) and the distributions are virtually identical (objective measure: *mean* = 25.9, *sd* = 4.8, *min.* = 15.8, *max.* = 57.8; self-report: *mean* = 25.4, *sd* = 4.5, *min.* = 16.0, *max.* = 56.2; n = 2'551). The same applies to the underlying measurements of body height and weight. This indicates a high external validity of the self-reported BMI.

Social variables

Finally, socio-economic, -demographic, and -cultural variables are surveyed in detail in KiGGS (educational background, parental job position, parental employment status, household size, sex, age, and migration background) as well as in NEMONIT (educational attainment, job position, employment status, partnership status, household size, sex, and age). All the variables included in the reported models of Tables 3 to 7 are described in Tables S1 to S4 of the Supplementary Information. As the Tables S1 and S3 reveal, diets are not a stable exposition, but underlie considerable variation within subjects over time. The same applies for all other time-varying characteristics shown in Tables S1 to S4.

All in all, the available KiGGS panel only enables the analysis of the social causation path, while all three paths of the nutrition cycle can be analyzed with the available NEMONIT panel data. It should be noted that a conceivable alternative for NEMONIT could be the German Health Interview and Examination Survey for Adults (DEGS). However, NEMONIT has more available observation waves than DEGS and in DEGS food intake is not measured identically in all waves.

Analytics

Concerning the analytics, fixed effects (FE) panel regression exploits the longitudinal data structure of panels as it only takes variations within the individuals' life courses into account. Thus, the FE estimator is unbiased in the presence of cross-sectional unobserved heterogeneity affecting both the observed covariates and the outcome [42,43]. If the strict exogeneity assumption holds, the FE regression adequately estimates unbiased causal effects of the covariates on the outcome[44]. For instance, this enables the identification of the effect of aging on dietary quality while controlling for birth cohort categories or whether an increase in dietary quality induces a decrease in BMI. However, it is still common practice in most of epidemiological prospective cohort studies not to regress changes in the outcome on changes in the covariates. This practice implicitly assumes that diets are a time-invariant exposition and remain stable over the individuals' life-courses[21]. This assumption is not empirically valid, as the KiGGS and NEMONIT data as well as Mertens et al. (2017)[17] demonstrate.

Nevertheless, standard FE models can only estimate effects for time-varying variables and do not allow the inclusion of time-invariant characteristics. The generalized linear mixed panel regression model (so-called hybrid model)[45,46] simultaneously estimates fixed, between, and random effects. Thus, the hybrid model enables the inclusion of both time-varying (e.g. income) and time-invariant (e.g. migration background) variables in the same model. As it is true for the FE model, the hybrid model's FE estimates are unbiased, if the strict exogeneity assumption holds[44]. Given the outlined advantages of the hybrid model, it is applied throughout the analyses. The specification of the hybrid model and an overview of the sensitivity analyses performed for all the reported regression results can be found in the supplement.

Results

Overall, the results indicate that gender, education, and age explain diet quality. Increases in the newly developed O-HEI-NVSII, and in nuts intake reduce BMI, while growing overall energy intake, lemonade, beer, meat, and meat products intake drive corpulence. In turn, developing obesity decreases socio-economic status. The results of the analyses for the social causation, biological causation, and health selection path are explained in more detail below.

Social causation

The social causation path could be analyzed for both the KiGGS and the NEMONIT panel. As introspection of Tables 3 and 4 consistently reveals, girls and women generally eat healthier (as measured by HuSKY and HEI) and have more diverse diets (as measured by the HFD Index) than boys and men. Moreover, children and adolescents with parents of high educational attainment have healthier and more diverse diets than peers with low educational background (see Table 3). The same applies to the educational attainment of adults. Furthermore, the dietary diversity of adults decreases with increasing age controlling for two-year birth cohort categories (see Table 4).

In addition to that, in KiGGS we observe an interaction effect between individual equivalence income and the average disposable county-level income on intake. With increasing individual income and high average disposable income in the county of residence (20’000-38’000 € per year) the intake of tolerated foods rises, while there is no relation for middle county income (15’000-<20’000 € per year) and a negative relation for low county income (10’000-<15’000 € per year) (see Figure 2). Interestingly, this interaction effect is not statistically significant for the intake of amply recommended food groups (see Figure 3).

For adults in NEMONIT, we found no effect of individual equivalence income on intake. Moreover, adults living in eastern Germany have higher HFD Index values than people living in the rest of Germany indicating higher food diversity in eastern Germany (see Table 4).

Table 3: KiGGS: Generalized Mixed Effects Regressions: Intake Quality

Model		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable		Amp.Rec.	Tolerated	Amp. Rec.	Tolerated	HuSKY	HFD Index	ED1	ED2
		g/d		kcal/d					
Equivalence Income (Unit: 1'000 €)	(F)	-12.01***	-7.33***	-5.81***	-14.42***	-0.03	0.15	-0.00	-0.00
		(3.06)	(1.83)	(1.43)	(3.09)	(0.14)	(0.15)	(0.00)	(0.00)
County-Level Disposable Income p.c.	(F)	-109.90***	-66.12***	-50.99***	-113.68***	-2.22	-0.40	-0.01	-0.02
		(30.30)	(18.82)	(14.21)	(27.94)	(1.40)	(1.49)	(0.02)	(0.03)
State Level Disposable Income p.c.	(F)	-117.01***	-21.86	-54.21***	-40.33+	-0.44	-1.66	0.02	0.02
		(32.31)	(20.48)	(15.98)	(20.89)	(1.40)	(1.47)	(0.02)	(0.03)
Interaction: Equivalence Income and County Level Disposable Income	(F)	57.74***	40.65***	26.79***	73.69***	0.53	-0.67	0.01	0.02
		(14.01)	(8.04)	(6.50)	(12.80)	(0.69)	(0.70)	(0.01)	(0.01)
Eastern Germany	(R)	30.17	20.18	28.87**	10.47	1.17	0.21	0.00	0.00
		(22.01)	(15.27)	(10.41)	(16.87)	(0.82)	(1.08)	(0.01)	(0.02)
Large City	(F)	53.35	-8.60	24.49	30.08	2.09	1.76	-0.03	-0.06
		(62.93)	(26.37)	(30.77)	(38.13)	(2.76)	(2.62)	(0.04)	(0.05)
Nr. of Firms p.c.	(F)	61.77	-7.86	34.71	24.41	3.57	4.80*	-0.01	0.04
		(48.97)	(29.48)	(23.25)	(44.31)	(2.35)	(2.45)	(0.03)	(0.05)
Interaction: Large City and Nr. of Firms p.c.	(F)	-65.76	69.24+	-27.59	13.19	-6.75+	-9.32**	0.03	0.03
		(72.68)	(35.59)	(34.93)	(58.65)	(3.46)	(3.39)	(0.05)	(0.07)
Firm Area 10-<20 ha/10000c	(F)	-3.83	8.68	-2.91	18.46	0.81	-0.76	0.01	0.00
		(24.99)	(15.10)	(11.99)	(18.96)	(1.09)	(1.24)	(0.02)	(0.02)
Firm Area 20-<80 ha/10000c	(F)	-18.77	12.64	-6.20	36.98	0.21	-0.75	0.01	-0.00
		(41.95)	(27.80)	(19.90)	(35.54)	(1.91)	(2.17)	(0.03)	(0.04)
Mother: Education	(B)	21.13***	-29.66***	8.39**	-14.51**	1.10***	2.23***	0.00	0.01
		(5.81)	(4.16)	(2.79)	(4.59)	(0.21)	(0.28)	(0.00)	(0.01)
Father: Education	(B)	21.28***	-20.70***	9.46***	-6.10	0.96***	1.77***	-0.00	0.01
		(5.57)	(3.98)	(2.66)	(4.63)	(0.20)	(0.27)	(0.00)	(0.01)
Parental Job Position	(F)	-10.65	-2.78	-4.87	-3.59	-0.32	-0.05	0.00	0.01
		(7.44)	(4.58)	(3.50)	(6.04)	(0.33)	(0.37)	(0.00)	(0.01)
Mother: Part-time Employed	(F)	-5.18	-9.44	-2.00	-9.60	-0.40	0.48	0.00	-0.00
		(17.82)	(10.27)	(8.40)	(13.58)	(0.79)	(0.86)	(0.01)	(0.02)
Mother: Full-time Employed	(F)	-7.83	-9.12	-1.21	-16.12	0.15	0.61	-0.01	-0.03
		(27.99)	(17.82)	(13.13)	(25.46)	(1.23)	(1.38)	(0.02)	(0.03)
Father: Part-time Employed	(F)	5.20	-10.78	6.67	19.72	1.94	2.45	0.02	-0.05
		(54.66)	(36.54)	(26.51)	(39.18)	(2.37)	(2.64)	(0.03)	(0.05)
Father: Full-time Employed	(F)	23.66	-54.11*	15.36	-55.16*	2.74	3.53+	-0.01	-0.07+
		(35.48)	(23.82)	(16.61)	(27.45)	(1.68)	(1.87)	(0.02)	(0.04)
Household Size	(F)	-14.60	5.46	-5.58	20.40	-1.00	-1.37+	0.02*	0.02
		(14.87)	(9.70)	(7.09)	(13.15)	(0.71)	(0.75)	(0.01)	(0.02)
Female	(R)	13.87	-67.81***	7.20+	-66.21***	3.64***	4.66***	-0.00	-0.08***
		(9.11)	(6.56)	(4.35)	(7.40)	(0.33)	(0.44)	(0.01)	(0.01)
Age	(F)	19.68	42.21*	8.17	34.01	2.34+	0.38	-0.01	0.03
		(33.12)	(18.77)	(15.59)	(26.13)	(1.36)	(1.54)	(0.02)	(0.03)
Age Squared	(F)	-3.83***	-2.46***	-1.75***	-4.17***	-0.03*	0.08***	-0.00***	-0.00
		(0.24)	(0.15)	(0.12)	(0.20)	(0.01)	(0.01)	(0.00)	(0.00)
Birth Cohort	(R)	-18.07*	-15.73**	-6.96+	-19.55**	-0.57+	0.33	0.01	-0.00
		(8.55)	(6.07)	(4.10)	(6.84)	(0.31)	(0.41)	(0.01)	(0.01)
Migration Background	(R)	8.84	-15.82+	7.04	6.59	0.78	0.66	0.01	-0.03*
		(13.80)	(9.50)	(6.62)	(12.12)	(0.52)	(0.66)	(0.01)	(0.01)
KiGGS 1	(F)	318.74	-55.08	134.96	14.37	-2.82	-2.30	-0.08	-0.37*
		(209.45)	(118.36)	(99.10)	(165.93)	(8.79)	(9.84)	(0.12)	(0.18)
Constant		841.74***	744.37***	363.96***	814.20***	48.14***	39.08***	0.69***	1.08***
		(148.61)	(100.98)	(71.11)	(114.61)	(5.17)	(6.87)	(0.09)	(0.14)
n x T		14173	14173	14173	14173	14173	14173	14173	14172
n		9299	9299	9299	9299	9299	9299	9299	9299
Log Pseudolikelihood		-108386.64	-103060.62	-97890.93	-105729.07	-62081.69	-65162.90	-3544.24	-10093.15

Note: + = $p<0.10$, * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$. Unstandardized regression coefficients with standard errors in brackets. (R) stands for random, (F) for fixed, and (B) for between effect. All standard errors are clustered by individual, and therefore robust with respect to heteroscedasticity and autocorrelation. data source: KiGGS Panel: Robert Koch-Institute. KiGGS = German Health Interview and Examination Survey for Children and Adolescents, amp. rec. = amply recommended food group intake (fruits, vegetables, and juice), tolerated = tolerated food group intake (sweets, and sweetened beverages), HuSKY = Healthy Nutrition Score for Kids and Youth, HFD Index = Healthy Food Diversity, ED1 = Energy density (non-beverages and caloric beverages), ED2 = Energy density (non-beverages only).

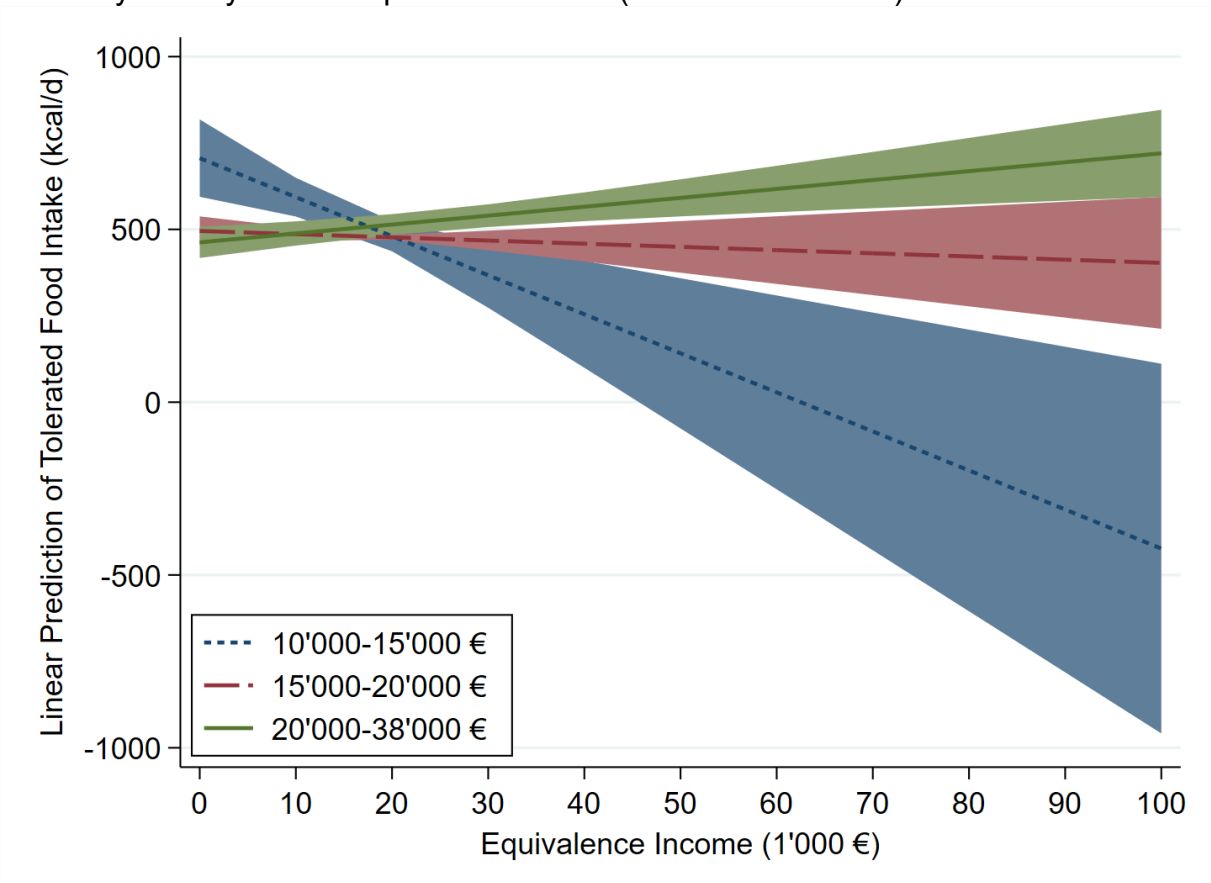
Table 4: NEMONIT: Generalized Mixed Effects Regressions: Intake Quality

Model		(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable		O-HEI-NVSII	HEI-MAC	HEI-EN	HFD Index	ED1	ED2
Equivalence Income	(F)	-0.05	-0.07	-0.02	0.05	-0.00	-0.00
(Unit: 1'000 €)		(0.03)	(0.05)	(0.06)	(0.05)	(0.00)	(0.00)
Eastern Germany	(R)	0.72	0.45	-0.57	2.45***	-0.00	-0.06**
		(0.48)	(0.59)	(0.64)	(0.68)	(0.02)	(0.02)
Education	(B)	0.60*	0.25	0.69*	1.26***	0.00	-0.01
		(0.24)	(0.31)	(0.32)	(0.37)	(0.01)	(0.01)
Job Position	(F)	-0.06	0.32	-0.10	0.19	0.00	0.00
		(0.17)	(0.26)	(0.30)	(0.29)	(0.01)	(0.01)
Employed	(F)	0.32	1.28	0.29	-1.21	0.02	0.00
		(0.63)	(1.02)	(1.13)	(1.16)	(0.02)	(0.02)
Single	(F)	-0.26	-0.43	2.65*	-2.38+	0.01	0.04
		(0.82)	(1.12)	(1.32)	(1.45)	(0.03)	(0.03)
Household Size	(F)	-0.15	-0.71	-0.37	-0.50	0.01	-0.01
		(0.36)	(0.48)	(0.58)	(0.61)	(0.01)	(0.01)
Female	(R)	2.82***	2.46***	1.32*	5.34***	0.00	-0.20***
		(0.39)	(0.50)	(0.51)	(0.64)	(0.01)	(0.01)
Age	(F)	-0.12	0.06	1.25	-2.70**	-0.02	-0.01
		(0.51)	(0.80)	(0.93)	(0.95)	(0.02)	(0.02)
Age Squared	(F)	-0.00	0.00	0.00	0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Birth Cohort	(R)	0.50	0.83	0.60	-0.95	0.02	-0.00
		(0.52)	(0.71)	(0.76)	(0.82)	(0.02)	(0.02)
Physical Exercise	(F)	0.01	0.02	0.02	0.05	0.00	-0.00
		(0.02)	(0.03)	(0.03)	(0.04)	(0.02)	(0.00)
Weekend	(B)	-0.18	-0.52	0.17	-0.51	0.01	-0.01
		(0.71)	(0.95)	(0.96)	(1.15)	(0.01)	(0.02)
Special Day	(B)	-0.57	0.47	-1.49+	-1.09	-0.01	0.05**
		(0.59)	(0.74)	(0.78)	(0.92)	(0.02)	(0.02)
Spring	(B)	0.66	-0.25	0.21	0.86	-0.05**	0.01
		(0.64)	(0.86)	(0.89)	(1.09)	(0.02)	(0.02)
Summer	(B)	0.79	-0.13	-0.89	1.59	0.02	0.00
		(0.65)	(0.88)	(0.94)	(1.09)	(0.02)	(0.02)
Fall	(B)	1.06	0.26	-0.88	1.00	-0.00	0.01
		(0.64)	(0.86)	(0.90)	(1.06)	(0.00)	(0.02)
NVSII (2005-07)	(F)	-0.58	3.55	8.19	-11.59*	-0.17+	-0.10
		(3.01)	(4.65)	(5.21)	(5.17)	(0.10)	(0.11)
2008/09	(F)	-0.36	1.54	5.63+	-4.16	-0.14*	-0.11
		(1.80)	(2.66)	(3.15)	(3.08)	(0.06)	(0.07)
2009/10	(F)	0.27	1.00	5.01*	-4.23+	-0.12**	-0.09+
		(1.30)	(2.00)	(2.28)	(2.25)	(0.04)	(0.05)
2010/11	(F)	0.10	1.45	1.25	-2.55	-0.08**	-0.08*
		(0.90)	(1.34)	(1.47)	(1.58)	(0.03)	(0.03)
Constant		42.19+	40.41	60.61+	95.26**	0.82	2.12**
		(22.12)	(30.08)	(32.02)	(34.74)	(0.77)	(0.80)
n x T		4114	4114	4114	4114	4114	4114
n		2243	2243	2243	2243	2243	2243
Log Pseudolikelihood		-14956.71	-16187.99	-16547.57	-16961.52	-941.29	-1223.53

Note: + = $p<0.10$, * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$. Unstandardized regression coefficients with standard errors in brackets. (R) stands for random, (F) for fixed, and (B) for between effect. All standard errors are clustered by individual, and therefore robust with respect to heteroscedasticity and autocorrelation. data source: NEMONIT: Max Rubner-Institute (2016)[25]. NEMONIT = German National Nutrition Monitoring, O-HEI-NVSII = Optimized Healthy Eating Index (HEI), HEI-MAC = HEI based on the macronutrient (MAC) intake recommendations, HEI-EN = HEI based on overall energy (EN) intake recommendations, HFD Index = Healthy Food Diversity, ED1 = Energy density (non-beverages and caloric beverages), ED2 = Energy density (non-beverages only).

Neither in KiGGS nor in NEMONIT the triple-A model can explain the energy density of diets as measured by ED1 including caloric beverages (see model 7 of Table 3 and model 5 of Table 4). As introspection of model 8 of Table 3 and model 6 of Table 4 reveals, girls and women have a lower average ED when computed excluding all beverages (ED2).

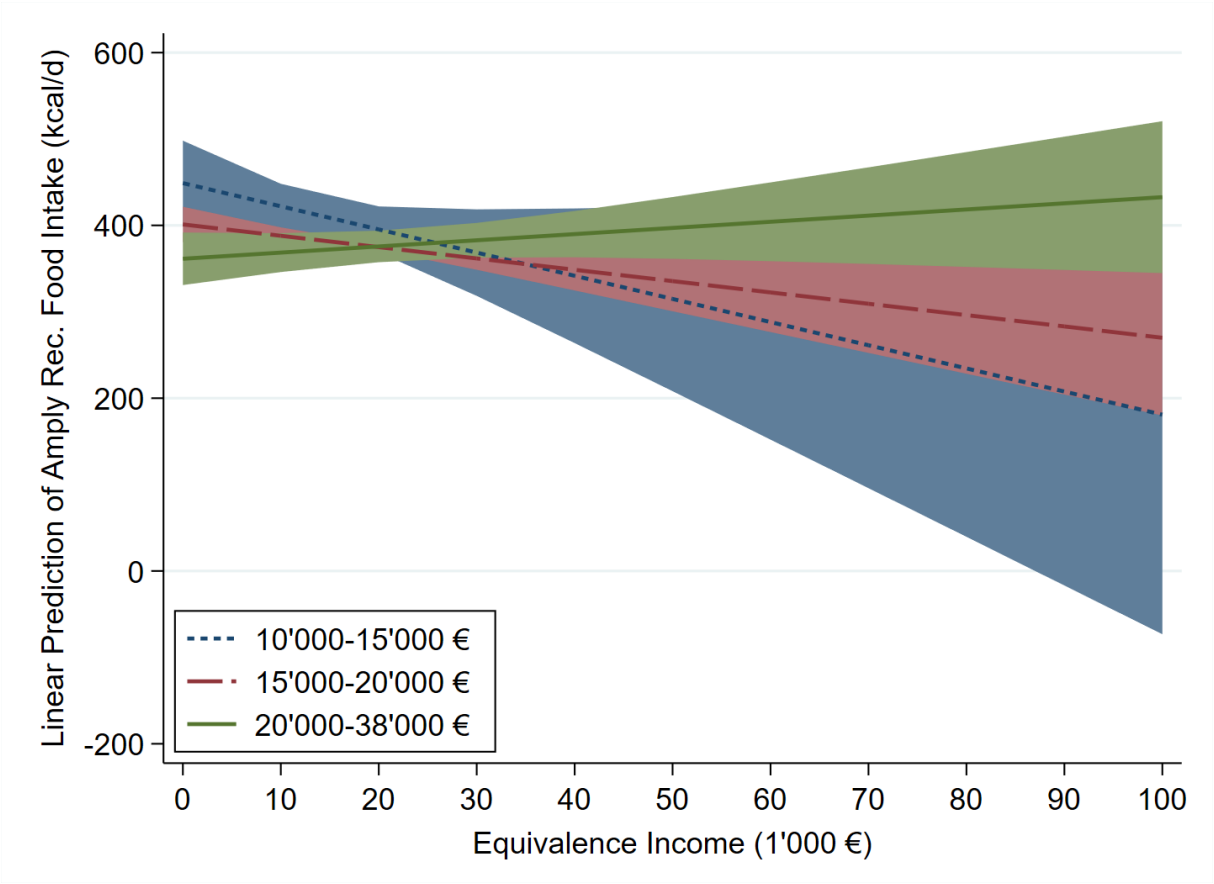
Figure 2: KiGGS: Predictive Margins of Equivalence Income for Tolerated Food Group Intake by County-Level Disposable Income (Model 4 of Table 3) with 95 % CIs



Note: The figure displays the interaction effect of equivalence income and county-level disposable income on tolerated food intake. Data sources: KiGGS Panel: Robert Koch-Institute. Country-level disposable income: Regional Database Germany (RDG) of the statistical offices of the confederation and the federal states.

All other variables included in the models do not substantially and consistently relate to healthy and diverse nutrition comparing the models using different indicators of diet quality (see Tables 3 and 4).

Figure 3: KiGGS: Predictive Margins of Equivalence Income for Amply Recommended Food Group Intake by County-Level Disposable Income (Model 3 of Table 3) with 95 % CIs



Note: The figure displays the interaction effect of equivalence income and county-level disposable income on amply recommended food intake. Data sources: KiGGS Panel: Robert Koch-Institute. Country-level disposable income: Regional Database Germany (RDG) of the statistical offices of the confederation and the federal states.

Biological causation

The analysis of the biological causation path (only possible for NEMONIT) reveals that increases in O-HEI-NVSII, and nuts intake reduce BMI, while growing overall energy intake, lemonade, beer, meat and meat products intake increase BMI (see Tables 5 and 6).

As models 1.1 and 1.2 of Table 5 show, a healthier diet decreases BMI. This indicates that the O-HEI-NVSII has some predictive validity. Nevertheless, the effect is very small. Per 10 index points healthier diet BMI decreases by 0.10. However, this is not the case for increases in intake diversity, as the HFD Index is not related to BMI in NEMONIT. Most surprisingly at first glance, growing energy density yields BMI losses net of overall energy intake, which has a relatively small positive effect

on BMI. The effect of ED including non-beverage food groups and caloric beverages is around twice as large as the effect of ED including non-beverage food groups only.

Models 2 and 3 of Table 5 reveal that independent from all other macronutrients only increases in alcohol consumption increase BMI substantially. As the food group-level analyses of the models 8 and 9 of Table 6 expose, this effect is only attributable to increases in beer consumption controlling for (sparkling) wine, spirits and all other food group-specific intake.

Further detailed influences on BMI can be obtained from the food group-specific models in Table 6: Nuts being very energy dense foods in combination with their unique composition of a variety of high-quality nutrients have various health promoting effects[47]. This study confirms clinical trials that suggest that an increase in nuts consumption yields a decrease in visceral adiposity[47]. Moreover, rising intake of lemonade, meat, and meat products ceteris paribus drive corpulence, with the effect of lemonade being six times as high as the effects of meat or meat products each. Changes in the intake of the other food groups do not affect BMI as it also applies for the categorization of food groups (see models 4 and 5 of Table 5). Finally, the models in Tables 5 and 6 also show that women have a lower BMI than men, and that age does not drive corpulence in the NEMONIT panel.

Table 5: NEMONIT: Generalized Mixed Effects Regressions of BMI: 1

Model		(1.1)	(1.2)	(2)	(3)	(4)	(5)
Dependent Variable		BMI	BMI	BMI	BMI	BMI	BMI
Intake in				100g/d	%kcal/d	100g/d	100kcal/d
O-HEI-NVSII	(F)	-0.01**	-0.01**				
		(0.00)	(0.00)				
HFD	(F)	-0.00	-0.00				
		(0.00)	(0.00)				
ED1 (non-beverages and caloric beverages)	(F)	-0.30***					
		(0.07)					
ED2 (non-beverages only)	(F)		-0.17*				
			(0.07)				
Overall Energy Intake	(F)	0.01*	0.01*				
		(0.00)	(0.00)				
Carbohydrates	(F)			0.05			
				(0.04)			
Fats	(F)			-0.01	-0.16		
				(0.10)	(0.30)		
Proteins	(F)			0.13	0.12		
				(0.12)	(0.60)		
Alcohol (Ethanol)	(F)			0.28*	1.00*		
				(0.11)	(0.42)		
Dietary Fibre	(F)			-0.29			
				(0.32)			
Amply Recommended	(F)					0.00	0.01
						(0.00)	(0.01)
Moderately Recommended	(F)					0.02	0.02*
						(0.01)	(0.01)
Sparsely Recommended	(F)					-0.14	-0.02
						(0.09)	(0.01)
Tolerated	(F)					0.02***	0.01
						(0.00)	(0.01)
Female	(R)	-1.67***	-1.71***	-1.65***	-1.61***	-1.64***	-1.66***
		(0.19)	(0.19)	(0.19)	(0.16)	(0.18)	(0.19)
Age	(F)	-0.04	-0.04	-0.04	-0.03	-0.05	-0.04
		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Birth Cohort	(R)	-0.01	-0.01	-0.01	-0.01	-0.02	-0.01
		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Physical Exercise	(F)	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Weekend	(B)	-0.07	-0.06	-0.03	0.08	0.20	0.04
		(0.36)	(0.36)	(0.35)	(0.35)	(0.36)	(0.35)
Special Day	(B)	0.03	0.05	0.29	0.35	0.05	0.17
		(0.29)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)
Spring	(B)	0.74*	0.75*	0.51	0.53	0.63+	0.66+
		(0.36)	(0.36)	(0.35)	(0.35)	(0.36)	(0.36)
Summer	(B)	0.73+	0.73+	0.61	0.54	0.57	0.68+
		(0.38)	(0.38)	(0.37)	(0.37)	(0.38)	(0.38)
Fall	(B)	0.19	0.19	0.14	0.11	0.04	0.13
		(0.34)	(0.34)	(0.33)	(0.33)	(0.34)	(0.33)
NVSII (2005-07)	(F)	-0.59+	-0.56+	-0.56+	-0.55+	-0.63*	-0.57+
		(0.31)	(0.31)	(0.31)	(0.31)	(0.31)	(0.31)
2008/09	(F)	-0.42*	-0.41*	-0.41*	-0.40*	-0.44*	-0.42*
		(0.18)	(0.18)	(0.18)	(0.19)	(0.18)	(0.18)
2009/10	(F)	-0.30*	-0.29*	-0.29*	-0.27*	-0.32*	-0.29*
		(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
2010/11	(F)	-0.22**	-0.22**	-0.22**	-0.21*	-0.23**	-0.22**
		(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Constant		27.48***	28.73***	24.47***	19.27***	23.01***	24.48***
		(2.84)	(2.98)	(2.78)	(2.83)	(2.77)	(2.78)
n x T		8446	8446	8446	8446	8446	8446
n		2582	2582	2582	2582	2582	2582
Log Pseudolikelihood		-17271.55	-17279.79	-17230.17	-17232.32	-17267.77	-17259.64

Note: + = $p<0.10$, * = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$. Unstandardized regression coefficients with standard errors in brackets. (R) stands for random, (F) for fixed, and (B) for between effect. All standard errors are clustered by individual, and therefore robust with respect to heteroscedasticity and autocorrelation. Data source: NEMONIT: Max Rubner-Institute (2016)[25]. BMI=Body mass index. NEMONIT=German National Nutrition Monitoring, O-HEI-NVSII=Optimized Healthy Eating Index (HEI), HFD Index=Healthy Food Diversity, ED1=Energy density (non-beverages + caloric beverages), ED2=Energy density (non-beverages).

Table 6: NEMONIT: Generalized Mixed Effects Regressions of BMI: 2

Model		(6)	(7)	(8)	(9)
Dependent Variable		BMI	BMI	BMI	BMI
Intake in		100g/d	100kcal/d	100g/d	100kcal/d
Bread, Cereals, Potatoes,	(F)	0.02	0.01		
Pasta, Rice		(0.02)	(0.01)		
Fruits	(F)	0.00	0.00		
		(0.01)	(0.02)		
Vegetables	(F)	0.01	0.05		
		(0.02)	(0.04)		
Milk (products)	(F)	-0.01	-0.01		
		(0.01)	(0.01)		
Meat, Meat Products, Fish, Eggs	(F)	0.07***	0.04***		
		(0.02)	(0.01)		
Edible Fats	(F)	-0.16^{+}	-0.02		
		(0.09)	(0.01)		
Confectionery, Snack Items	(F)	0.00	-0.00		
		(0.02)	(0.01)		
Water	(F)	0.01			
		(0.00)			
Coffee, Tea	(F)	-0.01^{+}	-0.67^{+}		
		(0.00)	(0.35)		
Juices	(F)	0.01	0.01		
		(0.01)	(0.02)		
Sweetened Soft Drinks	(F)	0.03**	0.13**		
		(0.01)	(0.05)		
Alcoholic Beverages	(F)	0.02***	0.03*		
		(0.01)	(0.01)		
Nuts, Seeds	(F)			-0.26*	-0.05*
				(0.13)	(0.02)
Meat	(F)			0.07*	0.03*
				(0.03)	(0.01)
Meat Products	(F)			0.08*	0.03*
				(0.03)	(0.01)
Fish	(F)			-0.02	-0.01
				(0.03)	(0.03)
Eggs	(F)			0.05	0.03
				(0.07)	(0.04)
Lemonade	(F)			0.04***	0.19**
				(0.01)	(0.06)
Beer	(F)			0.02***	0.05***
				(0.01)	(0.01)
(Sparkling) Wine	(F)			0.00	-0.00
				(0.01)	(0.02)
Spirits	(F)			-0.12	-0.07
				(0.10)	(0.06)
Female	(R)	-1.27***	-1.27***	-1.06***	-1.06***
		(0.19)	(0.20)	(0.20)	(0.20)
Age	(F)	-0.05	-0.04	-0.05	-0.04
		(0.05)	(0.05)	(0.05)	(0.05)
Birth Cohort	(R)	-0.03	-0.01	-0.03	-0.02
		(0.06)	(0.06)	(0.06)	(0.06)
Constant		22.56***	23.81***	22.25***	23.76***
		(2.73)	(2.74)	(2.73)	(2.74)
n x T		8446	8446	8446	8446
n		2582	2582	2582	2582
Log Pseudolikelihood		-17177.84	-17202.36	-17130.77	-17160.93

Note: $^{+}$ = p<0.10, * = p<0.05, ** = p<0.01, *** = p<0.001. Unstandardized regression coefficients with standard errors in brackets. (R) stands for random, (F) for fixed, and (B) for between effect. All standard errors are clustered by individual, and therefore robust with respect to heteroscedasticity and autocorrelation. All models additionally control for physical exercise, the observation waves, weekend, special day, and the seasons. Models (8) and (9) additionally control for all other food group specific intake. Data source: NEMONIT: Max Rubner-Institute (2016)[25]. BMI = Body mass index. NEMONIT = German National Nutrition Monitoring.

Health selection

In turn, the analysis of the health selection path (only for NEMONIT) reveals that developing obesity leads to losses in socio-economic status (see Table 7). Specifically, developing obesity decreases adults' job position as well as the associated prestige (model 2). However, developing obesity does not affect equivalence income (model 1). Reporting a good health condition is associated with a higher job position and women earn less and have lower job positions.

Table 7: NEMONIT: Generalized Mixed Effects Regressions for Germany: SES

Model		(1)	(2)
Dependent Variable		Equiv. Income	Job Position
Obesity (BMI ≥ 30)	(F)	-0.60	-0.22*
		(0.44)	(0.09)
Good Health Condition	(F)	0.06	0.12*
		(0.26)	(0.05)
Female	(R)	-1.52***	-0.50***
		(0.35)	(0.06)
Age	(F)	0.16	0.00
		(0.23)	(0.05)
Birth Cohort	(R)	0.68*	0.05
		(0.29)	(0.06)
Weekend	(B)	0.63	-0.04
		(0.75)	(0.12)
Special Day	(B)	2.67***	0.47***
		(0.64)	(0.10)
Spring	(B)	0.92	-0.05
		(0.76)	(0.13)
Summer	(B)	0.45	-0.14
		(0.75)	(0.13)
Fall	(B)	0.69	0.04
		(0.72)	(0.12)
NVSII (2005-07)	(F)	-1.49	-0.35
		(1.42)	(0.28)
2008/09	(F)	-1.80*	-0.03
		(0.83)	(0.17)
2009/10	(F)	-1.86**	-0.07
		(0.61)	(0.12)
2010/11	(F)	-1.08**	-0.07
		(0.39)	(0.08)
Constant		-12.89	1.37
		(12.15)	(2.34)
n x T		6891	6891
n		2373	2373
Log Pseudolikelihood		-23016.21	-11514.90

Note: $^{+}$ = p<0.10, * = p<0.05, ** = p<0.01, *** = p<0.001. Unstandardized regression coefficients with standard errors in brackets. (R) stands for random, (F) for fixed, and (B) for between effect. All standard errors are clustered by individual, and therefore robust with respect to heteroscedasticity and autocorrelation. The estimation of a model regressing education on obesity status was not possible because there is little within-variation in educational attainment over time. Data source: NEMONIT: Max Rubner-Institute (2016)[25]. SES = socio-economic status, BMI = body mass index. NEMONIT = German National Nutrition Monitoring.

Discussion and conclusion

For the first time, this study acquires comprehensive causal insight into the nutrition cycle by identifying the social causes of healthy diets and its impacts on obesity and socio-economic status. We analyze two German socio-epidemiological panel survey datasets (KiGGS and NEMONIT) covering an age range between 0 and 82 years and a timespan between 2003 and 2013. We exploit their longitudinal structure utilizing hybrid generalized linear mixed panel regression models. Generally, these models have the advantage of being able to cancel out the influence of changes in unobservables affecting both the observed covariates and the outcome. Thus, these models estimate unbiased causal effects (if strict exogeneity is given)[44]. In addition to that, hybrid models allow the inclusion of time-invariant variables.

Altogether, the results for the social causation path (a) indicate that gender, education, and age are able to explain healthy diets of German children, adolescents, and adults. Women and people with higher educational attainment/background eat healthier and have a more diverse diet. However, aging is associated with losses in dietary variety starting with adulthood. This finding is new, since former cross-sectional studies were only able to detect cohort effects. Our analysis of panel data can control for birth cohort and hence identify a genuine effect of aging. Furthermore, this disconfirms the assumption that in youth increasing peer influence and diminishing parental control over the offspring's diet makes it unhealthier and one-sided. This refers to the fact that this conception was based on negative cohort effects spotted in cross-sectional studies[16].

For income, we spotted an interaction effect between individual equivalence income and average county-level income on the intake of tolerated foods of children and adolescents. This finding might refer to differential availability structures between amply recommended and tolerated food groups in combination with individual social characteristics that lead to self-selection into rich and poor counties beyond the model at hand. Moreover, we found no direct effect of individual income on dietary quality and variety for both KiGGS and NEMONIT. In sum, these results indicate that affordability matters for healthy diets in KiGGS depending on contextual income only and not for NEMONIT. This expands former studies and generally conforms to the findings by Zagorsky and Smith (2017)[20] who discovered no effect of income on fast-food consumption frequency. Food availability was only possible to approximate

in KiGGS due to its relatively high number of subjects and sample points. None of the applied indicators (number of retail firms and firms in the food service industry p.c., factory area of trade and industry p.c.) have a substantial effect on intake in German children and adolescents. Also, moving into a large city with more than 100'000 inhabitants is not associated with changes in intake and diet quality disconfirming the findings of Zagorsky and Smith (2017)[20].

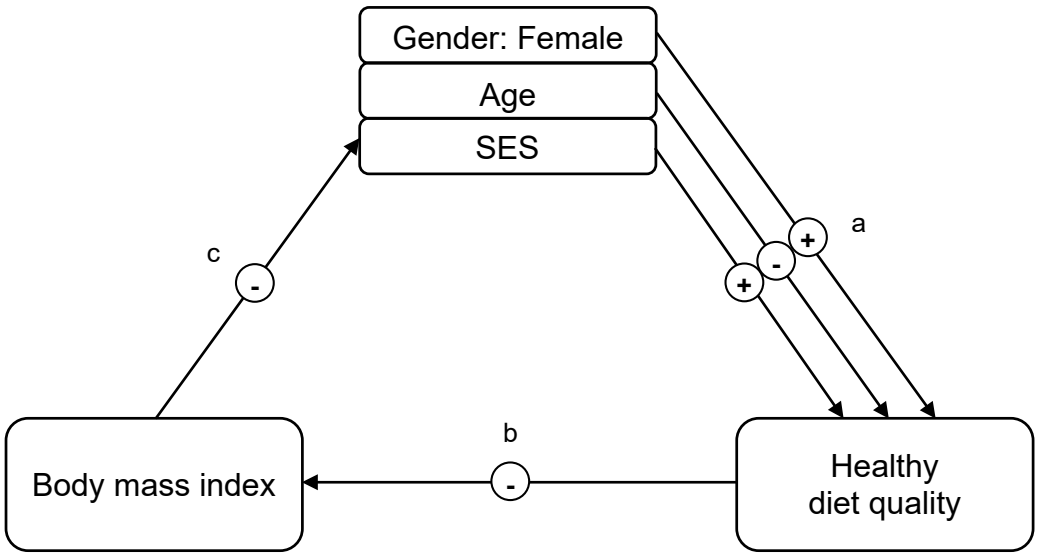
All other factors included in the models referring to the availability and accessibility of foods are not substantially linked to diets. The effects found in the KiGGS panel analysis differ from cross-sectional analyses for KiGGS that do find differences between subjects' diets with respect to employment status, migration background and age[16]. In sum, the applied triple-A model only had limited predictive power in the German context. Moreover, this study could not include further indicators of affordability (e.g. wealth), availability (e.g. number of food outlets within 1km distance from residence), and accessibility (e.g. nutritional health knowledge). Hence, future studies may focus on extending the model.

With regard to the biological causation path (b), for the first time, we thoroughly investigated the effects of changing dietary quality on BMI using a national socio-epidemiological panel study (NEMONIT). We found that increases in a newly developed Optimized Healthy Eating Index (O-HEI-NVSII) that presumably is less prone to assumptions on individual characteristics, slightly decreases BMI, indicating some predictive validity of O-HEI-NVSII. This also confirms the findings of previous studies[21-23]. Moreover and in accordance with clinical trials[47], growth in nuts intake reduces corpulence pointing to the health promoting properties of nuts. Growing lemonade, beer, meat, and meat products intake drive corpulence. The findings on the influence of the intake's average energy density are in contradiction to former research with cross-sectional data[40,48] and indicate that despite its intuitive appeal, the energy density hypothesis (more ED translates into more BMI) might be flawed because of a potential misconception of the construct of energy density that includes the water content of foods. A high ED does not necessarily mean less volume but also dryer food items[40,41]. Hence, high overall ED does not necessarily imply low satiation. What is more, there are food items like nuts that have a weight-lowering effect[47], as this study also demonstrates (see Table 6). In all, future research is vital for further validation of this finding.

As a whole, this panel analysis of the biological causation path identifies dietary predictors of corpulence. However, this study is not capable of eliminating the potential confounding influence of (epi)genetic disposition and composition of the gut's microbiome[49-53] and other genetic factors[54]. Thus, future research that takes account of these factors is essential for the validation of the findings.

In turn, the analysis of the health selection path (c) suggests that developing obesity leads to losses in socio-economic status. This might refer to both decreases in productivity[55] and weight stigma[56] that could explain the association between obesity and SES loss. Future research may delve deeper into analyzing this mechanism. Figure 4 summarizes the empirical results of our analysis of the nutrition cycle for Germany.

Figure 4: The nutrition cycle: empirical results for Germany.



Note: a = social causation, b = biological causation, c = health selection, + = positive effect, - = negative effect; The results for (a) are based on the panels of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) and of the German National Nutrition Monitoring (NEMONIT)[25]. The results for (b), and (c) are based on NEMONIT only. (a), (b), and (c) apply multiple generalized mixed effects regressions each.

Nevertheless, this analysis has some limitations: First, there is a methodological drawback in the KiGGS panel: the baseline survey used self-administered paper and pencil interviews, while in KiGGS 1 all interviews were taken out via telephone. Thus, potential measurement bias cannot be ruled out. However, the KiGGS and the NEMONIT analyses substantially lead to the same conclusions. Second, individual intake is surveyed via self-report. Social desirability and limited ability of

retrospection might lead to systematic biases especially concerning the descriptive statistics[57,58]. However, Willett (2013)[59] concludes that self-reported diet recall for periods of up to ten years can be reasonably accurate. Third, panel studies are probably subject to selective attrition. It is plausible that over time, the panel drop-out is selective, for example leading to an overrepresentation of relatively healthy subjects in the panel sample. The fact that individuals' health ratings are relatively stable over time might suggest that this is not much of a problem in NEMONIT. In addition, sample selection only leads to bias if both the covariates and the outcome are correlated with self-selection into the sample[60]. Fourth, the accuracy of the BMI in diagnosing obesity in terms of body fat percent is limited. Thence, future panel studies might utilize more valid – yet more costly – measurements of body fat percent like bioelectrical impedance analysis[61]. Fifth, the analysis is limited to a time span of about a decade covering only a small part of individuals' life spans. With socio-epidemiological panels getting longer, future studies will also be able to analyze the nutritional effects on mortality.

On balance, this study suggests that health promoting policies targeted at the reduction of obesity prevalence may be well-advised to focus on boys and men, people with low educational attainment level and background, as well as on the promotion of a healthy diet including nuts intake, and the limitation of lemonade, beer, meat, and meat products intake. The World Health Organization (WHO) and the Organisation for Economic Co-operation and Development (OECD) also urge the latter. Both the WHO and OECD recommend a stricter regulation for advertising unhealthy foods[62,1] and the WHO favors the implementation of a special tax of at least 20 % on sweetened beverages[63,64]. Moreover, the obesity-preventing effect of high levels of education suggests that the establishment of nutritional health equity would probably gain from the further development of setting-based health promotion. Especially advancing communal feeding at educational establishments from day-care centers to schools and in businesses' canteens is promising to this end[64,65]. Here, imposing mandatory catering standards could be effective[64]. Furthermore, the design of the decision architecture (i.e. 'nudging') is promising, as it is supposed to be more effective than traditional means of prevention[66-68]. Nudging measures like the prominent placing of healthy food items in canteens and grocery stores generally do not involve a social gradient in effectivity, as for example

it is the case with a traffic light-like labelling of foods[65,68-70]. Thus, nudging people into healthy diets may guarantee the primary aims of health promotion – individual autonomy of action and equality of opportunity. However, the challenge remains not only to make the healthy choices the easy choices but also to make them the preferred ones[65]. A promising avenue to attain this goal could be the placing of positive incentives, for instance via the implementation of consumer rewards for the purchase of healthy food items for example as part of loyalty cards of big grocery retailers[71].

Conclusions

In conclusion, for the first time, this study gains comprehensive causal insight into the nutrition cycle by identifying the social causes of healthy diets and its impacts on health and wealth in Germany. In all, the results indicate that girls and women, and people with high educational attainment level and background are less corpulent than boys and men, and people with low education. Healthy diets including nuts intake promote the reduction of obesity prevalence, while the intake of lemonade, beer, meat, and meat products counteract it. All told, this study advises further research to validate the findings and derive sound recommendations for political action using rigorous panel regression models and more accurate panel data. Many other longitudinal panel studies especially in Europe and the U.S. are still waiting for their longitudinal potentials to be exploited and made fruitful for causal inference.

References

1. OECD (2019) *The Heavy Burden of Obesity: The Economics of Prevention.* OECD Health Policy Studies. Paris: OECD Publishing.
2. The Global BMI Mortality Collaboration (2016) Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* **388**, 776-786.
3. Kivimäki M, Kuosma E, Ferrie JE *et al.* (2017) Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120813 adults from 16 cohort studies from the USA and Europe. *Lancet Public Health* **2**, e277-285.
4. Beyer F, Kharabian Masouleh S, Huntenburg JM *et al.* (2017) Higher Body mass index is associated with reduced default mode connectivity in older adults. *Hum Brain Mapp* **38**, 3502-3515.
5. Schwingshackl L, Schwedhelm C, Galbete C *et al.* (2017) Adherence to mediterranean diet and risk of cancer: An updated systematic review and meta-analysis. *Nutrients* **9**, 1063.
6. Zheng J, Guinter MA, Merchant AT *et al.* (2017) Dietary patterns and risk of pancreatic cancer: a systematic review. *Nutr Rev* **75**, 883-908.
7. GBD 2015 Risk Factors Collaborators (2016) Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1659-1724.
8. Konnopka A, Dobroschke A, Lehnert T *et al.* (2018) The costs of overweight and obesity: a systematic review. *Gesundheitswesen* **80**, 471-481.
9. Schwingshackl L, Bogensberger B. & Hoffmann G (2018) Diet quality as assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and health outcomes: an updated systematic review and meta-analysis of cohort studies. *J Acad Nutr Diet* **118**, 74-100.
10. Cawley J & Ruhm CJ (2012) The economics of risky health behaviors. In *Handbook of Health Economics*, pp. 95-199 [M Pauly, T McGuire & PP Barros, editors]. Oxford: North Holland.
11. Muff C (2009) *Soziale Ungleichheiten im Ernährungsverhalten. Theoretische Hintergründe und empirische Befunde.* Berlin/Wien: LIT Verlag.
12. Kröger H, Pakpahan E & Hoffmann R (2015) What causes health inequality? A systematic review on the relative importance of social causation and health selection. *Eur J Public Health* **25**, 951-960.
13. Fekete C & Weyers S (2016) Social inequalities in nutrition. Evidence, causes and interventions. *Bundesgesundheitsblatt* **59**, 197-205.
14. Reisch L & Gwozdz W (2011) Von der „Macht der Defaults“ und vom „sanften Stupsen“: Verhaltensökonomische Erkenntnisse als Impulse für eine effektive Ernährungspolitik. In *Die Zukunft auf dem Tisch. Analysen, Trends und Perspektiven der Ernährung von morgen*, pp. 323-336 [G Hirschfelder, A Ploeger and G Schönberger, editors]. Wiesbaden: VS Verlag für Sozialwissenschaften.
15. Story M, Neumark-Sztainer D & French S (2002) Individual and environmental influences on adolescent eating behaviors. *J Am Diet Assoc* **102**, 40-51.
16. Mader S, Rubach M & Schaecke W (2017) The impact of social factors on the nutrition of German children and adolescents. *Ernahrungs Umschau* **64**, 96-100.

17. Mertens E, Clarys P, Mullie P *et al.* (2017) Stability of physical activity, fitness components and diet quality indices. *Eur J of Clin Nutr* **71**, 519-524.
18. Solbak NM, Xu J-Y, Ven JE *et al.* (2017) Diet quality is associated with reduced incidence of cancer and self-reported chronic disease: Observations from Alberta's Tomorrow Project. *Prev Med* **101,** 178-187.
19. Rummo PE, Guilkey DK, Ng SW *et al.* (2017) Understanding bias in relationships between the food environment and diet quality: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *J Epidemiol Commun H* **71**, 1185-1190.
20. Zagorsky JL & Smith PK (2017) The association between socioeconomic status and adult fast-food consumption in the U.S.. *Econ Hum Biol* **27**, 12-25.
21. Mertens E, Deforche B, Mullie P *et al.* (2015) Longitudinal study on the association between three dietary indices, anthropometric parameters and blood lipids. *Nutr Metab* **12**, 47.
22. Fung TT, Pan A, Hou T, *et al.* (2015) Long-Term Change in Diet Quality Is Associated with Body Weight Change in Men and Women. *J Nutr* **145**, 1850-1856.
23. Cespedes Feliciano EM, Tinker L, Manson JAE, et al. (2016) Change in Dietary Patterns and Change in Waist Circumference and DXA Trunk Fat Among Postmenopausal Women. *Obesity* **24**, 2176-2184.
24. Michaud P-C & van Soest A (2008) Health and wealth of elderly couples: causality tests using dynamic panel data models. *J Health Econ* **27**, 1312-1325.
25. Max Rubner-Institute (2016) *Dataset of the German National Nutrition Monitoring (NEMONIT)*. Karlsruhe, Germany.
26. Truthmann J, Mensink GBM & Richter A (2011) Relative validation of the KiGGS Food Frequency Questionnaire among adolescents in Germany. *Nutr J* **10**, 133.
27. Mauz E, Gösswald A, Kamtsiuris P *et al.* (2017) New data for action. Data collection for KiGGS Wave 2 has been completed. *Journal of Health Monitoring* **2**, 2-27.
28. Gose M, Krems C, Heuer T *et al.* (2016) Trends in food consumption and nutrient intake in Germany between 2006 and 2012: results of the German National Nutrition Monitoring (NEMONIT). *Brit J Nutr* **115**, 1498-1507.
29. HafenCity University Hamburg, University Regensburg (2013) *Qualifizierte Nahversorgung im Lebensmitteleinzelhandel.* https://www.hcu-hamburg.de/fileadmin/documents/Professoren_und_Mitarbeiter/Projektentwicklung__-management/Forschung/Qualifizierte_Nahversorgung_HCU-IREBS_Endbericht__Langfassung__Maerz_2013.pdf (accessed December 2018).
30. Mensink GB & Burger M: (2004) Was isst du? Ein Verzehrshäufigkeitsfragebogen für Kinder und Jugendliche. (What do you eat? A food frequency questionnaire for children and adolescents). *Bundesgesundheitsblatt* **47**, 219-226.
31. Kleiser C, Mensink GBM, Scheidt-Nave C *et al.* (2009) HuSKY: a healthy nutrition score based on food intake of children and adolescents in Germany. *Brit J Nutr* **102**, 610-618.
32. Alexy U, Clausen K & Kersting M (2008) Die Ernährung gesunder Kinder und Jugendlicher nach dem Konzept der Optimierten Mischkost. *Ernahrungs Umschau* **3/08**, 168-177.

33. Kersting M, Alexy U & Clausen K (2005) Using the Concept of Food Based Dietary Guidelines to Develop an Optimized Mixed Diet (OMD) for German Children and Adolescents. *J Pediatr Gastr Nutr* **40**, 301-308.
34. Drescher LS, Thiele S & Mensink GBM (2007) A New Index to Measure Healthy Food Diversity Better Reflects a Healthy Diet Than Traditional Measures. *J Nutr* **137**, 647-651.
35. Slimani N, Deharveng G, Charrondiere RU *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Meth Prog Bio* **58,** 251–266.
36. Oberritter H, Schäbethal K, von Ruesten A et al. (2013) The DGE-nutrition circle – presentation and basis of the food-related recommendations from the German Nutrition Society (DGE). *Ernahrungs Umschau* **60,** 24–29.
37. Gedrich K & Karg G (2001) Welche Ernährung wäre für die Bevölkerung optimal? In *Ernährungsziele unserer Gesellschaft: die Beiträge der Ernährungsverhaltenswissenschaft. 22. Wissenschaftliche Jahrestagung der Arbeitsgemeinschaft Ernährungsverhalten e. V.*, pp. 19-38 [U Oltersdorf and K Gedrich, editors]. Karlsruhe: Bundesforschungsanstalt für Ernährung.
38. DGE – Deutsche Gesellschaft für Ernährung e.V., Österreichische Gesellschaft für Ernährung, Schweizerische Gesellschaft für Ernährungsforschung, Schweizerische Vereinigung für Ernährung (2013) *Referenzwerte für die Nährstoffzufuhr.* Frankfurt: Umschau-Verlag.
39. DGE – German Nutrition Society (2015) New Reference Values for Energy Intake. *Ann Nutr Metab* **66**, 219-223.
40. Drewnowski A, Almiron-Roig E, Marmonier C *et al.* (2004) Dietary energy density and body weight: Is there a relationship? *Nutr Rev* **62**, 403-413.
41. Shrapnel B (2010) Is energy density a useful concept for understanding and preventing obesity? *Nutr Diet* **67,** 281-286.
42. Brüderl J & Ludwig V (2015) Fixed-effects panel regression. In *The SAGE handbook of regression analysis and causal inference*, pp. 327–358 [H Best and C Wolf, editors]. London: SAGE.
43. Wooldridge J (2010) *Econometric analysis of cross-section and panel data.* Cambridge: MIT Press.
44. Vaisey S & Miles A (2017) What you can – and can't – do with three-wave panel data. *Sociol Method Res* **46**, 44-67.
45. Allison PD (2009) *Fixed Effects Regression Models.* Thousand Oaks, CA: Sage.
46. Schunck R & Perales F (2017) Within- and Between-Cluster Effects in Generalized Linear Mixed Models: A Discussion of Approaches and the xthybrid Command. *Stata J* **17,** 89-115.
47. Ros E (2010) Health benefits of nuts consumption. *Nutrients* **2**, 652-682.
48. Karl JP & Roberts SB (2014) Energy density, energy intake, and body weight regulation in adults. *Adv Nutr* **5**, 835-850.
49. Ley RE, Turnbaugh PJ, Klein S *et al.* (2006) Human gut microbes associated with obesity. *Nature* **444**, 1022-1023.
50. Menni C, Jackson MA, Pallister T *et al.* (2017) Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. *Int J Obesity* **41**, 1099-1105.
51. Ott B, Skurk T, Hastreiter L *et al.* (2017) Effect of caloric restriction on gut permeability, inflammation markers, and fecal microbiota in obese women. *Sci Rep-UK* **7**, 11955.

52. Selber-Hnatiw S, Rukundo B, Ahmadi M *et al.* (2017) Human gut microbiota: Toward an ecology of disease. *Front Microbiol* **8**, 1265.
53. Sommer F, Moltzau Anderson J, Bharti R *et al.* (2017) The resilience of the intestinal microbiota influences health and disease. *Nat Rev Microbiol* **15**, 630-638.
54. Cohen-Cline H, Lau R, Moudon AV *et al.* (2015) Associations between fast-food consumption and body mass index: a cross-sectional study in adult twins. *Twin Res Hum Genet* **18**, 375-382.
55. Goettler A, Grosse A & Sonntag D (2017) Productivity loss due to overweight and obesity: a systematic review of indirect costs. *BMJ Open* **7**, e014632.
56. Brewis AA (2014) Stigma and the perpetuation of obesity. *Soc Sci Med* **118**, 152-158.
57. Hill RJ & Davies PSW (2001) The validity of self-reported energy intake as determined using the doubly labelled water method. *Brit J Nutr* **85**, 415–430.
58. Westerterp KR & Goris AH (2002) Validity of the assessment of dietary intake: problems of misreporting. *Curr Opin Clin Nutr* **5**, 489-493.
59. Willett W (2013) *Nutritional Epidemiology.* New York: Oxford University Press.
60. Elwert F & Winship C (2014) Endogenous selection bias: The problem of conditioning on a collider variable. *Annu Rev Sociol* **40**, 31-53.
61. Romero-Corral A, Somers VK, Sierra-Johnson J *et al.* (2008) Accuracy of body mass index in diagnosing obesity in the adult general population. *Int J Obesity* **32**, 959-966.
62. WHO – World Health Organization (2014) *European Food and Nutrition Action Plan 2015-2020.* Regional Committee for Europe. Copenhagen. EUR/RC64/14.
63. Spiller A, Zühlsdorf A & Nitzko S (2017a) Instruments of Nutrition Policy. A research overview – Part 1. *Ernahrungs Umschau* **64**, M146-M153.
64. Spiller A, Zühlsdorf A & Nitzko S (2017b) Instruments of Nutrition Policy. A research overview – Part 2. *Ernahrungs Umschau* **64**, M204-M210.
65. Hawkes C, Smith TG, Jewell J *et al.* (2015) Smart food policies for obesity prevention. *Lancet* **385**, 2410-2421.
66. Kenning P & Reisch L (2013) Alternativen zum Informationsparadigma der Verbraucherpolitik. 1. Verbraucherforum an der Zeppelin Universität, Friedrichshafen, 24. Januar 2013. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **8**, 227-253.
67. Thaler RH & Sunstein CR (2009) *Nudge. Improving decisions about health, wealth and happiness.* London: Penguin Books.
68. Thorgeirsson T & Kawachi I (2013) Behavioral Economics. Merging Psychology and Economics for Lifestyle Interventions. *Am J Prev Med* **44**, 185-189.
69. Kelly B, Hughes C, Chapman K *et al.* (2009) Consumer testing of the acceptability and effectiveness of front-of-pack food labeling systems for the Australian grocery market. *Health Promot Int* **24**, 120-129.
70. Glanz K, Bader MDM & Iyer S (2012) Retail grocery store marketing strategies and obesity. An integrative review. *Am J Prev Med* **42**, 503-512.
71. Willett W, Rockström J, Loken B, Springmann M, Lang T, et al. (2019) Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. The Lancet. DOI: 10.1016/S0140-6736(18)31788-4.

Supplementary Material of "Healthy nutrition in Germany: A survey analysis of social causes, obesity and socio-economic status"

Model specification

The hybrid model[45,46] can be written as

$$y_{it} = (\boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_i)\boldsymbol{\beta_W} + \bar{\boldsymbol{x}}_i\boldsymbol{\beta_B} + \boldsymbol{c}_i\boldsymbol{\beta_R} + \boldsymbol{z}_t\boldsymbol{\gamma} + u_i + \varepsilon_{it}. \quad (1)$$

y_{it} denotes the dependent variable of individual *i* at time *t*. $\boldsymbol{x}_{it}$ stands for the vector of all exogenous time-varying variables for *i* at time *t*, and $\bar{\boldsymbol{x}}_i$ for the mean of the whole observation period. $\boldsymbol{c}_i$ variables only vary between clusters but are time-invariant. The model also comprises a vector of dummy variables (***z***) for every point in time, which controls period effects for all individuals. u_i denotes the random intercept, and an individual's time varying stochastic error term is represented by ε_{it}.

Sensitivity analyses

All the reported regression results were tested for robustness: First, all models were recalculated by performing fixed effects regression with individual slopes and constants to allow for heterogeneous growth (FEIS)[42,43,72] only for NEMONIT, as FEIS regression is only executable for at least three-waves panel data. Second, all models were rerun excluding influential cases from the regression dropping individuals with Cook's D>1. The robustness of standard errors was investigated via non-parametric bootstrapping. None of these checks had any substantial influence on the estimates. Moreover, all parameters were tested for linearity including penalized splines FE models[73]. Further sensitivity checks comprise the implementation of different operationalizations of different constructs depending on data availability. For KiGGS, physical exercise, single parent status, and psychological problems were available as further control variables. These were not included in the main analyses, because of comparatively low number of person years (see Table S2 in the Supplement). However, none of these variations affected the reported results in any substantial way. In addition, the robustness of all estimates with respect to model specification was assessed using the procedure suggested by Young and Holsteen (2017)[74]. Furthermore, the potential influence of omitted variables was examined using the method suggested by Frank (2000)[75]. Also, all these checks detected no fundamental

deviations from the reported results. All models comprise the analysis of complete cases (CC). Being aware of the potential benefits of multiple imputation (MI; more efficient and less biased estimates), we reran the models with multiply imputed data. The results of CC and MI analyses are the same. This is due to the small differences in the number of missings between outcomes and independent variables (see Tables S1-S4). In this case, the potential benefits of MI are negligible, since CC performs equally well[76]. All the analyses were conducted using the statistical software package STATA 15.1.

Table S1: KiGGS: Variable description of all Outcomes

Variable	Description	Mean/ Share	within ($\bar{x}_i$) sd	min.	max.	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) sd	min.	max.	*N* (*n* x *T*)	*n*
Amply Recommended Food Groups (g/d)	Summed intake (g/d and kcal/d) of amply recommended (fruits, vegetables, and juice) and tolerated (sweets, and sweetened beverages) food group intake. Food group specific intake in g/d is calculated from intake frequencies and amount/portion per intake (ref. 77: 19f) and kcal/d as calculated from g/d and food group specific consumption[78] weighted average energy densities[79-81].	797.6	328.3	-1416.7	3011.9	497.7	0	4571.4	26722	16416
Amply Recommended Food Groups (kcal/d)		370.1	158.2	-545.9	1286.0	238.6	0	1993.3	26722	16416
Tolerated Food Groups (g/d)		368.1	227.6	-1175.8	1912.1	428.3	0	3881.5	26657	16390
Tolerated Food Groups (kcal/d)		521.3	323.8	-2901.7	3944.3	530.4	0	8280.6	26657	16390
HuSKY	Healthy Nutrition Score for Kids and Youth (HuSKY) following Kleiser et al.[31] based on intake of fruits, vegetables, juice, sweets, and sweetened beverages in g/d, standardized between 0 and 100.	49.1	12.8	.4	97.8	17.4	0	100	26058	16136
HFD Index	Healthy Food Diversity (HFD) Index following Drescher et al.[34] based on intake in g/d and relative recommended intake ratio of amply recommended to tolerated food groups according to the Optimized Mixed Diet[32,33]; standardized between 0 and 100.	46.5	13.7	-1.0	93.4	23.7	0	100	26056	16135
ED1 (non-beverages and caloric beverages)	Average energy density (ED) of caloric food intake excluding non-caloric beverages and drinking water in kcal/g[40].	.79	.20	-.50	2.09	.29	.28	3.21	26056	16135
ED2 (non-beverages only)	Average energy density of non-beverage food intake excluding all beverages in kcal/g[40].	1.12	.33	-.27	2.52	.49	.27	3.21	26296	16228

Note: data source: KiGGS Panel: Robert Koch-Institute. KiGGS = German Health Interview and Examination Survey for Children and Adolescents.

Table S2: KiGGS: Variable description of all Covariates

Variable	Description	Mean/ Share	within ($\bar{x}_i$) sd	within ($\bar{x}_i$) min.	within ($\bar{x}_i$) max.	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) sd	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) min.	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) max.	*N* (*n* x *T*)	*n*
Equivalence Income (micro level)	Calculated according to German Council for Economic Experts[82] as ratio of household net income and square root of household size. Unit: EUR per year.	16609.2	4114.0	-89876.4	123094.7	7736.4	1341.6	138564.1	24311	16820
County Level Disposable Income per Capita (p.c.)	Average disposable income p.c. of county of residence appended to the KiGGS data from the Regional Database Germany (RDG) of the statistical offices of the confederation and the federal states (1=10000-<15000, 2=15000-<20000, 3=20000-<38000 EUR per year). It represents the mean of the years 2003-2006, and 2009-2012 according to the 2 waves of KiGGS.	2.1	.4	1.6	2.6	.4	1	3	32572	16286
State Level Disposable Income p.c.	Average disposable income p.c. of federal state of residence appended to the KiGGS data from the RDG. Unit: EUR per year.	18050.4	952.4	13886.4	22214.4	1935.4	14131.8	21739.9	29632	17640
Eastern Germany	Binary, 1, if place of residence in Eastern Germany.	.33	.05	-.17	.83	.47	0	1	29632	17640
Large City	Binary, 1, if place of residence with more than 100000 inhabitants.	.23	.10	-.27	.73	.41	0	1	29632	17640
Retail Firms and Hospitality Industry p.c.	Number of retail firms and firms in the food service industry per capita of county of residence (data source: RDG). Binary, 1, if 15<=nr./1000c<28.	.10	.10	-.40	.60	.28	0	1	33792	16896
Factory Area p. c.	Factory area of trade and industry (excl. mining area) of county of residence (data source: RDG; 1=0-<10, 2=10-<20, 3=20-<80 ha/10000c).	1.6	.2	1.1	2.1	.7	1	3	33792	16896
Mother: Education	Five-point scale according to ISCED97 of the mother's/father's educational attainment (school+vocation).	3.6	.3	1.6	5.6	1.0	1	5	24710	16700
Father: Education		3.8	.3	1.8	5.8	1.0	1	5	23096	16038
Parental Job Position	Seven-point scale of the socio-economic status of the job position of the household's main breadwinner[77].	3.4	.5	.4	6.3	1.3	1.1	7	24903	17185
Mother: Employment Status	1=unemployed, 2=part-time employed, 3=full-time employed;	1.8	.3	.8	2.8	.7	1	3	25521	17285
Father: Employment Status		2.8	.2	1.8	3.8	.6	1	3	23908	16604
Single Parent*	Binary, 1, if single parent responsible for child.	.13	.05	-.37	.63	.33	0	1	15175	14522
Household Size	Nr. of people living in the same household.	3.9	.3	1.9	5.9	.8	1	5	25665	17294
Sex: female	Binary.	.50	0	.50	.50	.50	0	1	29632	17640
Age	In complete years.	10.8	2.7	6.8	14.8	5.1	0	24	29632	17640
Birth Cohort	In 2 year groups (1985-1986=1,…,2005-2006=11)	6.2	0	6.2	6.2	2.5	1	11	23984	11992
Migration Background	Binary.	.22	0	.22	.22	.41	0	1	35120	17560
Physical Exercise*	Binary, 1, if physical activity level is high.	.68	.09	.18	1.18	.46	0	1	14816	14205
Psychological Problems*	Binary (ref. 83: 110).	.10	.14	-.40	.60	.27	0	1	22938	16473

Note: variables with '*' are only included in the robustness analyses, as N is comparatively low. Data source: KiGGS Panel: Robert Koch-Institute. KiGGS = German Health Interview and Examination Survey for Children and Adolescents.

Table S3: NEMONIT: Description of all nutrition variables

Variable	Description	Mean/ Share	within ($\bar{x}_i$) sd	within ($\bar{x}_i$) min.	within ($\bar{x}_i$) max.	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) sd	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) min.	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) max.	*N* (*n* x *T*)	*n*
Amply Recommended Food Groups (g/d) (1)	Sum of food group specific intake in g/d and kcal/d as calculated from g/d and food group specific consumption[78] weighted average energy densities[79-81]. (1) = fruits + vegetables + cereal products + potatoes + coffee + tea (+ water (g/d only)). (2) = dairy products + meat products + fish + eggs. (3) = edible fats. (4) = alcoholic beverages + sweetened soft drinks + juices + sweets + snacks.	2422.6	462.6	-561.6	5170.0	732.4	483.7	5523.2	9462	2610
Amply Recommended Food Groups (kcal/d) (1)		743.9	167.6	-199.2	2287.3	224.6	126.9	2380.4	9462	2610
Moderately Recommended Food Groups (g/d) (2)		375.0	128.1	-538.5	1973.9	164.1	13	1343.5	9462	2610
Moderately Recommended Food Groups (kcal/d) (2)		512.6	156.3	-230.6	1920.7	190.3	29.3	2131.9	9492	2610
Sparsely Recommended Food Groups (g/d) (3)		27.6	14.4	-77.2	150.7	17.5	0	160.4	9462	2610
Sparsely Recommended Food Groups (kcal/d) (3)		196.1	95.9	-488.7	879.6	123.0	0	952.4	9462	2610
Tolerated Food Groups (g/d) (4)		642.6	301.7	-1578.4	3396.6	461.9	1	3158.0	9454	2610
Tolerated Food Groups (kcal/d) (4)		663.3	256.0	-634.6	2472.0	310.4	2.2	2490.3	9454	2610
Overall Intake (g/d)	Sum of intake (1) + (2) + (3) + (4) in g/d and kcal/d.	3467.8	498.4	917.7	6211.5	741.6	1569.2	8046.6	9454	2610
Overall Intake (kcal/d) (5)		2115.7	367.0	62.3	4452.0	571.2	460.2	5488.7	9454	2610
Carbohydrate Intake (g/d)	Macronutrient intake in g/d and energy percent.	233.8	46.6	18.5	570.0	71.3	61.0	697.2	9462	2610
Carbohydrate Intake (% of kcal/d)		45.9	5.7	21.7	78.0	6.9	21.9	70.4	9462	2610
Fat Intake (g/d)		81.8	20.7	-41.7	226.7	28.4	15.6	228.7	9462	2610
Fat Intake (% of kcal/d)		35.6	5.2	10.1	59.2	5.6	14.5	58.1	9462	2610
Protein Intake (g/d)		73.4	16.0	-4.5	191.8	21.0	26.2	188.4	9462	2610
Protein Intake (% of kcal/d)		14.6	2.4	4.4	29.5	2.5	7	28.7	9462	2610
Alcohol Intake (g ethanol/d)		12.3	10.5	-92.4	143.1	15.2	0	138.0	9462	2610
Alcohol Intake (% of kcal/d)		3.9	3.0	-15.2	37.5	4.6	0	35.4	9462	2610
Dietary Fibre Intake (g/d)		20.9	5.1	-18.9	60.7	7.2	3.5	71.7	9462	2610
O-HEI-NVSII (0-100)	Optimized HEI-NVSII according to relative food group and gender specific intake recommendations after Gedrich and Karg (2001)[37].	58.2	6.3	30.9	81.9	7.7	35.5	84.7	9454	2610
HEI-MAC (0-100)	HEI based on the macronutrient (MAC) intake recommendations of DGE et al. (2013)[38].	67.4	9.0	31.0	101.7	9.9	38.9	99.0	9462	2610
HEI-EN (0-100)	HEI based on overall energy (EN) intake recommendations of DGE (2015)[39].	84.3	10.0	26.4	125.9	9.8	28.5	100	9454	2610
HFD Index (0-100)	Healthy Food Diversity (HFD) Index following Drescher et al. (2007)[34] (based on intake in g/d).	73.8	10.2	3.7	123.8	13.0	12.3	97.0	9462	2610
ED1 (non-beverages and caloric beverages)	Average energy density (ED) of caloric food intake excluding non-caloric beverages and drinking water in kcal/g[40].	1.34	0.20	0.23	2.50	0.26	0.61	2.54	9454	2610
ED2 (non-beverages only)	Average energy density of non-beverage food intake excluding all beverages in kcal/g[40].	1.68	0.22	0.73	2.90	0.29	0.78	2.76	9454	2610

Note: The intake of all food groups is based on the mean of the two 24h-recall measurements. Data source: NEMONIT: Max Rubner-Institute (2016)[25]. NEMONIT = German National Nutrition Monitoring.

Table S4: NEMONIT: Description of all health, socio-economic, socio-demographic, and control variables

Variable	Description	Mean/ Share	within ($\bar{x}_i$) sd	min.	max.	between($x_{it} - \bar{x}_i + \bar{\bar{x}}$) sd	min.	max.	*N* (*n* x *T*)	*n*
BMI	Body mass index (BMI).	25.6	.9	16.7	34.7	4.4	16.4	62.7	9769	2610
Obese	Binary, 1, if BMI ≥ 30.	.14	.14	-.66	.94	.32	0	1	9769	2610
Good Health Condition	Binary, 1, if self-rated health condition good or very good.	.81	.24	.01	1.61	.32	0	1	9774	2610
Equivalence Income (micro level)	Calculated according to German Council for Economic Experts (ref. 82: 262) as ratio of household net income and square root of household size. Unit: EUR per year.	21063.3	4224.0	-10616.8	48963.3	8599.4	1697.1	66000	8286	2504
Eastern Germany	Binary, 1, if place of residence in Eastern Germany.	.18	.01	-.32	.98	.38	0	1	9783	2610
Education	Five-point scale according to ISCED97 of the educational attainment (school+vocation).	3.9	.4	2.3	5.9	.9	1	5	9447	2590
Job Position	Seven-point scale of the socio-economic prestige of the job position[77].	4.5	.9	.1	8.5	1.5	1	7	9217	2587
Employment Status	Binary, 1, if employed.	.58	.21	-.22	1.41	.45	0	1	9593	2608
Single	Binary, 1, if single	.38	.18	-.42	1.18	.41	0	1	6094	2609
Household Size	Nr. of people living in the same household.	2.6	.5	-.6	6.6	1.1	1	7	9773	2610
Sex: female	Binary.	.57	0	.57	.57	.49	0	1	13050	2610
Age	In complete years.	51.6	2.2	46.9	55.6	15.9	15.5	81.7	9783	2610
Birth Cohort	In 2 year groups (1927-1928=0, …, 1991-1992=32).	15.0	.3	14.2	15.8	7.9	0	32	9783	2610
Physical Exercise	In hours per week of light sports activities. Middle sports activities are doubled and heavy sports activities are quadrupled according to caloric influence[84].	8.3	6.6	-62.2	82.3	9.7	0	158.5	9648	2610
Season	Season of interviews (Spring=1, Summer=2, Fall=3, Winter=4). Included as season dummies into the models (reference: Winter).	2.5	.9	.1	4.9	.6	1	4	9783	2610
Weekend	Binary, 1, if interviews on weekends (Saturday and Sunday).	.18	.33	-.62	.98	.22	0	1	9462	2610
Special Day	Binary, 1, if interviews on a special day (holiday, vacation, illness, shift work).	.20	.32	-.60	1.00	.26	0	1	8572	2586

Note: data source: NEMONIT: Max Rubner-Institute (2016)[25]. NEMONIT = German National Nutrition Monitoring.

Additional References

72. Polachek SW & Kim MK (1994) Panel estimates of the gender earnings gap: individual-specific intercept and individual-specific slope models. *J Econometrics* **61**, 23–42.
73. Ruppert D, Wand MP & Carroll RJ (2003) *Semiparametric Regression*. Cambridge, UK: Cambridge University Press.
74. Young C & Holsteen K (2017) Model Uncertainty and Robustness: A Computational Framework for Multimodel Analysis. *Sociol Method Res* **46**, 3-40.
75. Frank KA (2000) Impact of a confounding variable on a regression coefficient. *Sociol Method Res* **29**, 147-194.
76. Karahalios A, Baglietto L, Lee KJ *et al.* (2013) The impact of missing data on analyses of a time-dependent exposure in a longitudinal cohort: a simulation study. *Emerging Themes in Epidemiology* **10**, 6.
77. RKI – Robert Koch-Institut (Ed.) (2008) *Public Use File KiGGS, Kinder- und Jugendgesundheitssurvey 2003-2006. Dokumentation des Datensatzes.* Berlin.
78. BMELV – Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2008) *Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten der Bundesrepublik Deutschland.* Bremerhaven: Wirtschaftsverlag NW.
79. Souci SW, Fachmann W & Kraut H (2000) *Die Zusammensetzung der Lebensmittel. Nährwert-Tabellen.* Stuttgart: medpharm Scientific Publishers.
80. BMELV – Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (2014) *Bundeslebensmittelschlüssel BLS 3.01*.
81. Nutri-Science GmbH (2014) *Prodi 6.2 Expert*.
82. SBGE – Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (Ed.) (2000) *Jahresgutachten 2000/01*. Berlin: Deutscher Bundestag.
83. Woerner W, Becker A, Friedrich C *et al.* (2002) Normierung und Evaluation der deutschen Elternversion des Strengths and Difficulties Questionnaire (SDQ): Ergebnisse einer repräsentativen Felderhebung. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* **30**, 105-112.
84. Ainsworth BE, Haskell WL, Whitt MC *et al.* (2000) Compendium of Physical Activities: an update of activity codes and MET intensities. *Med Sci Sport Exer* **32**, 498-516.