AN EMPIRICAL STUDY FOR THE ACCEPTANCE OF ORIGINAL NUDGES AND HYPERNUDGES

Yukari Yamazaki

Seikei University (Japan)

yyamazak@econ.seikei.ac.jp

ABSTRACT

While nudges have been paid attention to, they have also been criticized by several studies especially for their ethicality. Sunstein (2016, 2018) considered several aspects of nudges and claimed that if nudges guarantee decision-makers autonomy and dignity as well as transparency of choice architecture, they are extremely beneficial for individuals and society and are therefore not unethical.

In the recent past, the neologism 'hypernudge' has drawn attention, arousing much controversy. It is said that a hypernudge is a 'kind of' nudge utilizing artificial intelligence (AI) or machine learning. While the ethicality of nudges is likely to be certified, keeping autonomy, dignity, and transparency in mind, it is certainly arduous to warrant these three conditions in hypernudges driven by AI artefacts. That is, the acceptance of the original nudges, which have become popular and been adhered to, are likely to be distinct from that of hypernudges.

To address the issue, this study tested the acceptance of both interventions and revealed that the acceptance of the original nudges differed from that of hypernudges, and that the latter was less acceptable than the former. Notably, in hypernudges, while individuals tended to accept the less flexible and forbidden intervention, they rejected the ones that utilized their children's personal data. However, neither typical nor common features were confirmed that could identify the acceptance level of hypernudges, such as categories of interventions, individual sociodemographic factors, political attitudes, and mobile phone usage histories. The findings from this study suggest a kind of alert for spreading hypernudges that utilize AI-driven artefacts in the future.

KEYWORDS: AI-driven artefacts, original nudge, hypernudge, acceptance.

1. INTRODUCTION

The paradigm word, nudge (Thaler and Sunstein, 2008) and its strategy has been attracting people in various fields. One breakthrough attempt is changing the law around organ donation (Max and Keira's law) from the year 2020 in the UK. All adults in England will be considered organ donors when they die unless they have recorded a decision not to donate or are in one of the excluded groups (BBC News, 2019). Nowadays, other countries such as Denmark and the Netherlands will also change to or consider adopting an opt-out organ donation system, a nudging technique. This paternalistic strategic intervention tries either presenting people in a

more salient or impressive light, or making them the easier or default option, rather than enforcing restrictions or drawing out people's rational behaviour. Such a selection system is called the choice architecture. Notably, a nudge promotes people's choice and behaviour and is assumed to benefit target individuals and society as a whole; therefore, it should never be applied to marketing or particular profit pursuing activities.

While the fact that a nudge steering peoples' behaviour in desirable directions through milder choice interventions has drawn attention, it has also received blistering critiques of ethicality (e.g., Goodwin, 2012; O'Neil, 2011), diminishing human wisdom (Furedi, 2011), troubles and pitfalls (Bovens, 2009), and manipulations (Wilkinson, 2012). In response to these misunderstood critiques, C. R. Sunstein, one of the advocates of the nudge, has discussed the validity and considered the benefit and ethicality of nudges (Sunstein, 2015, 2016). According to his consideration, there are neither neutral ways to present options nor can choices be made in a vacuum, and one cannot avoid the choice architectures that influence choice in many ways. It might be easy to promote purchasing by altering the presentation order of alternatives and attributes, ease of picking them up, selection of defaults, and naming just a few of the design options available. Therefore, it is essential to choose alternatives while paying attention to the structures and effects of the choice architectures.

He also said, 'When nudges are fully transparent and subject to public scrutiny, a convincing ethical objection is less likely to be available'. In addition, he also stated that, 'if people have not consented to them; such nudges can undermine autonomy and dignity' (p.1). Furthermore, Sunstein (2018) insisted that 'Nudges always respect, and often promote human agency; because nudges insist on preserving freedom of choice, they do not put excessive trust in government; nudges are generally transparent rather than covert or forms of manipulation' (p.1). Indeed, it has been already examined that nudges utilized the defaults setting to be transparent and yet effective (Bruns et al., 2018). According to the above considerations, two of the prominent elements in ethical nudges should be transparency and autonomy. In other words, maintaining transparency and decision-makers' autonomy in nudging must be recognized as ethical and beneficial.

Currently, the neologism 'hypernudge' has highlighted and aroused much controversy. It is thought of as a 'kind of' nudge utilizing big data, personal data, AI algorithms, deep learning, and so on. While the ethicality of nudges would be certified, keeping autonomy, dignity, and transparency in mind, it is certainly arduous to warrant these three conditions in hypernudges driven by AI artefacts. That is, the acceptance of the original nudges, which have become popular and been applied to, would be distinct from that of hypernudges. While empirical evidence regarding the acceptance of the original nudges has appeared, the ones of hypernudges have remained unproven.

In this study, the comparison with acceptance of the original nudge and hypernudge is examined. It is found that while hypernudges are less acceptable than the original nudges, the representative features showing which hypernudges are more acceptable than others are still veiled. The findings are discussed, and conclusions are drawn at the end.

2. HYPERNUDGE

Recently, artificial intelligence (AI)/machine learning (ML) has drawn attention among mass media and academic fields not only because of their attractive, tremendous, and hyper functions

as well as efficiency and effectiveness, but also because of their ethicality and riskiness. The IEEE, for example, has taken the ethicality of AI designing, utilizing, and prevalence as a serious problem and given an alert for AI systems as nudging tools. (IEEE, 2018). In the section Affective Computing of the 2nd draft version of *Ethically Aligned Design*, the following six recommendations have been pointed out: 1) systematic analysis for ethical design of AI systems before deployment, 2) showing the types, effects, and purposes of nudges towards users, 3) analysing the possibility of infantilization of those who were nudged by AI, 4) making default settings opt-in, 5) giving additional protection for vulnerable users who do not pay attention to informed consent; and 6) keeping transparency and accountability. These are consistent with the certified requirements of nudges mentioned above, guaranteeing transparency and autonomy.

Nowadays, nudges through AI/ML-driven new technologies are coined as 'hypernudges' (Yeung, 2017) or 'digital nudges' (Weinmann et al., 2016). While it is being gradually known that AI/ML systems have beneficial traits, there are several specific features as hypernudges. Some of them are, for example, self-tracking of past behaviour (in some cases) without getting the agreement for utilizing it, presenting immediate feedback based on self-tracking and big data as recommendations for each user, making some judgment on behalf of human autonomy, and steering people to use AI artefacts repeatedly (e.g., Google navigation system). Based on prior studies that had paid attention to hypernudge, this study considers several differences between the original nudges and hypernudges (Table 1).These typical features with AI/ML-driven hypernudges might make users blind and depend too much on them. Therefore, the manipulative aspects of data-driven personalized communication, big data utilization, and behavioural targeting in the online realm have been regarded as problems (e.g., Lanzing, 2018).

Table 1. Comparison of the original nudges and hypernudges.

	Original nudge	**Hypernudge**
Purpose	Steering people towards a better direction to nudge softly and mildly utilizing human judgmental tendencies.	Designing and programming the choice architecture utilizing bigdata, personalized information, and computer algorithms.
Methods	• Presenting visible, available, noticeable information or options. • Using the default setting to do something automatically, omit doing something, or avoid forgetting. • Emphasizing (the possibility of) suffering losses, gaining incentives, and so on. • Utilizing social influence such as normative message, showing others behaviour, or giving approval.	• Presenting desirable options such as recommendations, alerts, advertisements, or notices that are suited to each individual based on the person's behaviour history.
Examples	• Labelling healthy food packages, setting fruits and vegetables nearer to a person in the buffet counter, providing information about the amount of sodium or sugar, serving food with a small plate, and so on. • Web shopping sites recommend storing of customers' credit card numbers to avoid entering these details again at the next shopping visit.	• Recommending purchasing healthier, lower sodium and sugar food based on the user's purchase history and health condition in the following shopping visit. • Web shopping sites recommend which credit cards the customer should use in this shopping.

Specific features	Giving a general message, not using specific customized data.	• Giving customized recommendation or predictive message based on personal information or bigdata. • Vague request for consent for using personal information.
	One to a few nudges. (A few people are nudged.)	One to many nudges. (It is possible to nudge many people at once. The presented information is different for individuals.)
	Transient and does not influence the next or other choices.	Continuous and repetitive and influences the next or other choices, as well as the possibility to change the choice architecture.
	No particular feedback, generally.	Immediate and continuous feedback based on personal data and bigdata.

In another argument, however, whereas various services by AI/ML, such as vehicle navigation systems, positional information on the digital map, recommendations based on purchase history, personalized chatting with bots, various apps (e.g., health care, saving and investing money with FinTech, and smart home with IoT), and so on have spread among people recently, it is hard to ascertain these new nudges as the original ethical nudges, whose validity has been discussed in Sunstein (2015, 2016) because of several serious reasons (Yamazaki, 2019). On the one hand, the following three factors have been pointed out as the hampering factors of autonomy. The first is decreasing users' motivation, responsibility, and morality. The second is indulging in of bad habits because of attractive recommendations from AI/ML artefacts. The third is a semiconscious repetition of habitual behaviour against their real will (meta preference). On the other hand, transparency is infringed by the complicated algorithm, dynamic feedback that confuses users' preferences, unpredictability, and complete unavailability of users' informed consent. Because of the lack of autonomy and transparency, it is difficult to certify hypernudges as original nudges at the moment.

3. PRIOR STUDIES

In the past years, several studies have surveyed the acceptance, trustworthiness, and consensus for various types of interventions in various countries under different contexts. These surveys have shown that, on the one hand, citizens in various countries perceived that nudges were being inconsistent with their interests or values of most choosers; on the other hand, they generally tended to approve of almost all nudges.

As for the difference of countries (nationality), three prior studies (Reisch and Sustein, 2016; Sunstein et al., 2018) investigated the approval of the same 15 interventions for 15 nations. According to their survey, the approval rates of many western democratic countries (Australia, Canada, French, Germany, Italy, Russia, the UK, and the US) were similar (around 68-75%). In contrast, Brazil, South Africa, China, and South Korea showed overwhelmingly high approval rates for all nudges (around 80%), while rates for Denmark, Hungary, and Japan were extremely low (around 60%). They could find neither the reasons nor countries' specific features, therefore, requiring further examination.

The fields of health and safety nudges would be approved for people and the levels of acceptance of nudging techniques depended on the countries of the participants as well as the depth, types, contexts, and prosociality of nudges. As in other empirical studies on the acceptance of nudges, in the fields of medicine and health, choice and public policies (e.g.,

Diepeveen et al., 2013; Felsen et al., 2013; Junghans et al., 2015, 2016; Jung and Mellers, 2016) showed similar results; whereas almost all subjects generally approved of nudges, they did not accept the nudges inconsistent with their preferences, improper nudges such as political and religious favouritism and perceived manipulation.

Other causes of different acceptances are the types and contexts of nudging. Felsen et al. (2013) tried a decisional enhancement nudging program that contained two types of nudging questions on five scenarios. One is overt and conscious nudging which the decision maker is aware of and can consciously process, while another is covert and unconscious nudging such as subconsciously decreasing hunger, which in some situations, can be related to people's autonomy. The five scenarios promote healthier eating, prudent online purchasing, encouraging exercising, investing in retirement, and improving productivity. The results showed that conscious nudges are more acceptable than subconscious nudge processes except for improving productivity scenarios. The subconscious nudges might infringe upon individual autonomy, while covertly trying to protect their own autonomous decision making.

Jung and Meller (2016) examined the effects of types of nudges (automatic and unconscious vs. effortful and conscious), individual dispositions (e.g., level of empathy, conservative, desire for control, reactance, and individualist vs. communitarian) and benefit from nudges (i.e., nudges for societal vs. personal). They divided several nudges into System 1 (automatic and unconscious), and 2 (effortful and conscious). Whereas the automatic enrolment for something such as retirement savings or medical coverage plans and default settings were classified as System 1, providing of reminders, alerts, or messages were classified as System 2. The results showed that the effect of the nudge type on support was significant. The effortful nudges were significantly more accepted than the automatic ones. In addition, the level of support for nudges was not significantly affected by benefits from nudges but was affected by participants' individual dispositions.

Although the acceptance of the original nudge as mentioned above has been gradually positive, empirical studies that have paid attention to the hypernudge utilized AI/ML artefacts have been scarce.

4. RESEARCH QUESTIONS AND METHODS

The aim of this research is to examine whether the acceptance of interventions differs from the original nudge to hypernudges. In particular, because hypernudges driven by AI artefacts would keep neither transparency nor autonomy for decision-makers, the high level of acceptance can come under doubt. Namely, while the original nudges have revealed almost acceptable results among almost all people, hypernudges that are new, neither transparent nor autonomous might have lower acceptance than the original nudges. In addition, according to several prior studies, the acceptance levels differed in the categories of intervention as well as several demographic factors. In sum, hypotheses concerning people's acceptance of interventions are as follows:

H1: The acceptance rate might differ in the original nudge and the hypernudge.

H2: The original nudge might be more acceptable than a hypernudge.

H3: The categories of interventions might be related to the peoples' acceptance rate of both the original nudge and hypernudge. Deeper and subconscious interventions might be less approved.

H4: Sociodemographic variables, political attitudes, and years of mobile use might be related to the acceptance rate of both the original nudge and the hypernudge. Whereas age might positively relate to acceptance, the years of mobile use might have a negative relationship with it.

To examine these hypotheses, this study compares and enlarges the results of prior studies. The same 15 survey interventions in the background literature (e.g., Sunstein et al., 2018) were selected, with one new intervention on security protection added for both the original nudges and hypernudges. The summary of each content is shown in Table 2. An example of a question on hypernudge is 'Online food shopping sites are required to show good or bad effects of each food on each user's health condition (such as decreasing body fat, body pressure, no effects, and so on) based on AI recommendation'. Because participants might not be familiar with and imagine for each different AI-driven service, general and uniform expression about AI services, 'artificial intelligence will be expected to present various recommendations using your personalized information such as your age, health and medical history, sleeping time, career, status of income and assets, purchase histories, and mobile phone use history in the future' was included at the beginning of the nudging questions. Participants were asked to answer 'agree' or 'disagree' to 16 questions on the social systems that promoted health and enriched society as well as to comment on nudges by AI.

Table 2. Categories of 16 interventions.

No.	Summary of contents	Depth	Type	Context	Prosociality	Percentage of agreement※
1	Showing calorie labels in restaurants' menu.	Mandatory information disclosure	Conscious	Health	Personal	79.40
2	Showing food bad effects for health.	Mandatory information disclosure	Conscious	Health	Personal	80.87
3	Enrolling green energy suppliers automatically, possible to opt out.	Mandatory default	Subconscious	Ecology	Social	73.47
4	Asking to be organ donors in obtaining driver's licence.	Mandatory default	Subconscious	Charity	Social	62.53
5	Placing healthy foods at prominent visible places in grocery stores.	Mandatory default	Subconscious	Health	Personal	73.00
6	An education campaign to reduce distracted driving.	Government Campaign	Conscious	Traffic safety	Social	87.67
7	An education campaign for promoting healthier choice for parents to reduce childhood obesity.	Government Campaign	Conscious	Health	Social	89.73
8	Providing prohibited subliminal advertisements in theatres to discourage smoking and overeating.	Non-nudge (Forced)	Subconscious	Health	Personal	52.00
9	Charging a specific amount with offset opt out option for carbon emission.	Mandatory default	Subconscious	Ecology	Social	44.00
10	Labelling unhealthy food making notice it is harmful.	Mandatory information disclosure	Conscious	Health	Personal	83.67
11	Asking to donate the Red Cross refund automatically, possible to opt out.	Mandatory default	Subconscious	Charity	Social	40.33
12	Requiring movie theatres to provide public education messages to discourage smoking and overeating.	Mandatory information disclosure	Subconscious	Health	Personal	67.13

13	Requiring large electricity providers to make people enrol in green energy suppliers automatically, possible to opt out.	Mandatory default	Subconscious	Ecology	Social	72.27
14	Keeping cashier areas in supermarkets chains free of unhealthy foods to halt obesity.	Mandatory choice architecture	Subconscious	Health	Personal	62.86
15	Requiring public institutions to have meat-free day per week.	Mandatory choice architecture	Conscious	Health	Personal	55.14
16	Installing security software automatically to avoid viruses and hackers, possible to opt out.	Mandatory default	Subconscious	Information security	Personal	

※The average rates of agreement that prior studies examined.

In addition, this study refers to the categories of interventions as laid out in prior studies (Jung and Meller, 2016; Felsen, 2013; Sunstein et al., 2018) (Table 2). It is categorized into five levels of depth: the shallowest is campaign Nos. 6 and 7); second, mandatory information disclosure (Nos. 1, 2, 10, and 12); third, mandatory default (Nos. 3, 4, 5, 9, 11, 13, and 16), fourth, mandatory choice architecture (Nos. 14 and 15), and the deepest is forced (No. 8). There are also two types (conscious vs. unconscious), five contexts (health, ecology, charity, traffic safety, and security), and two prosocialities (for society or personal) of interventions.

Participants in this study were recruited from university students in Japan and the Japanese consulting company *Kiccoe Survey* and students attending university in Tokyo. A total of 1,192 participants were asked 16 questions, half of which were on original nudges and the other half on hypernudges. The original nudge had *n*=596 and the hypernudge had *n*=596. The percentage of males was 54.6%. The mean age was 36.64: *SD*=18.2, range=13–87 years, under 20=332 (27.85%), the 20s=209 (17.53%), the 30s=140 (11.74), the 40s=175 (14.68%), the 50s=162 (13.59%), the 60s=118 (9.90%), the 70s=49 (4.11%), and the 80s=7 (0.59%). The participants also gave responses on their years of education (*M*=14.98 years), years of mobile phone use (never=134 or 11.24%, under a year=44 or 3.78%, under 5 years=374 or 31.38%, under 10 years=515 or 43.21%, and over 10 years=124 or 10.40%), and political attitude (ruling=308 or 25.84%, opposition=173 or 14.51%, and non-partisan=711 or 59.65%).

The first was designed to compare the acceptance levels of each original nudge and the hypernudges. The second was designed to test the differences among the categories of interventions, followed by examining the effects of participants' individual dispositions as the third.

5. RESULTS

5.1 Overall acceptance rates

The results indicated that overall, the original nudges were more accepted than hypernudges (Table 3), and the acceptance level of the original nudge and hypernudge differed in the categories of interventions (Table 4). In line with prior studies on the original nudges, more than half of the original nudges (10 out of 16) were significantly more acceptable. On the other hand, as new notions, less than half of the hypernudges (6 out of 16) were significantly acceptable. In addition, eight interventions (Nos. 1, 2, 3, 4, 6, 7, 9, and 12) of the original nudges, half of the total, found significantly higher acceptance than hypernudges. These results indicate that people would not be more receptive to AI-driven interventions and suggest H1 and H2 are

supported. One possible reason for these results is that using their personal information or behavioural history might have been considered creepy and untrustworthy; although hypernudges are customised for the individuals, their acceptance level was lower than the original nudge. The results would suggest that AI artefacts designers, service providers, choice architectures as well as users, should be careful of hypernudges, not a little.

Table 3 Significant difference of acceptances within each original nudge and hypernudge.

Original nudges (*n*=596)

No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
Agree	514**	395**	444**	349*	272	407**	469**	249	***Means of original nudges*** Agree 347.68 Disagree 248.31 % of agree 58.34
Disagree	82	201	152	247	324*※	188	127	347**	
% of agree	86.24	66.28	74.50	58.56	45.64	68.29	78.69	41.78	
No.	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	
Agree	327*	341**	281	394**	294	215	183	429**	
Disagree	269	255	315	202	302	381**	413**	167	
% of agree	54.87	57.21	47.15	66.11	49.33	36.07	30.70	71.98	

Hypernudges (*n*=596)

No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
Agree	306	327*	352**	315	269	281	199	328*	***Means of hypernudges*** Agree 300.19 Disagree 295.75 % of agree 50.38
Disagree	290	269	244	281	327*	315	397**	268	
% of agree	51.34	54.87	59.06	52.85	45.13	47.15	33.39	55.03	
No.	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	
Agree	285	379**	291	295	329*	227	162	459**	
Disagree	311	217	305	301	267	369**	434**	137	
% of agree	47.82	63.59	48.83	49.50	55.20	38.09	27.18	77.01	

Chi-square value significant at alpha * $p < 0.05$

** $p < 0.01$

※The numbers with underlines suggest significantly more unacceptable results.

Table 4. Cross Tabulation with each nudge pairs.

	O1	H1	O2	H2	O3	H3	O4	H4	O5	H5	O6	H6	O7	H7	O8	H8
Agree	**62.7**※	37.3	**54.7**	45.3	**55.8**	44.2	**52.6**	47.4	50.3	49.7	**59.2**	40.8	**70.2**	29.8	43.2	**56.8**
Disagree	22.0	78.0	42.8	57.2	38.4	61.6	46.8	53.2	49.8	50.2	37.5	62.5	24.2	75.8	56.4	43.6
$\chi^2(1)$	169.062**		16.243**		32.007**		3.930*		.030		54.576**		248.254**		20.964**	
	O9	H9	O10	H10	O11	H11	O12	H12	O13	H13	O14	H14	O15	H15	O16	H16
Agree	**53.6**	46.4	47.4	**52.6**	49.1	50.9	**57.2**	42.8	47.2	**52.8**	48.6	51.4	53.0	47.0	48.3	**51.7**
Disagree	46.2	53.8	54.0	46.0	50.8	49.2	40.2	59.8	53.1	46.9	50.8	49.2	48.8	51.2	54.9	45.1
$\chi^2(1)$	6.500*		5.065*		.336		33.710**		4.119*		.518		1.799		3.974*	

Chi-square value significant at alpha * $p < 0.05$

** $p < 0.01$

※The significant higher acceptances are shown with bold letters

It might be easy to imagine that because No. 8 (subliminal advertisement) were, as Sunstein et al. (2018) also mentioned, not a nudge, there were significantly higher disagreements in the original nudge. In addition, No. 5 (placing visible healthy food), 14 (avoiding unhealthy food),

and 15 (requiring meat-free day per a week) were significantly lower acceptance in both the original nudges and hypernudges because they were strongly obtrusive for people.

Unlike prior studies, the highest accepted intervention in the original nudges was the 'calorie label' (No. 1, 86.24%), with the next being 'education campaign for childhood obesity' which was in highest agreement in prior studies. The lowest accepted interventions in the original nudges, on the other hand, were 'requiring a meat-free day per week for health' (No. 15, 30.7%), which was likewise the lowest in hypernudges (27.18%) and relatively lower acceptance in prior studies of hypernudges. This might have been considered as excessive interference in diet. One of the higher accepted interventions in both nudges was 'information security' (No. 16) that might have been the most familiar matter for the largest number of youths in this study. Overall, the acceptance tendency of the original nudges shown in this study is similar to that of Sunstein et al. (2018), where the Japanese consensus was remarkably lower than in other countries and in averages.

As for comparison with the original nudges and hypernudges for acceptance percentages (Table 4), the number of acceptance of the eight original nudges (Nos. 1, 2, 3, 4, 6, 7, 9, and 12) were exceed that of hypernudges, as mentioned, while the one of the four hypernudges (Nos. 8, 10, 13, and 16) exceed that of original nudges, exceptionally. The other four interventions (Nos. 5, 11, 14, and 15) were not significantly different between the original and the hypernudge. Surprisingly, whereas No. 8 (subliminal advertisement) intervention is significantly higher in the original nudges, it is significantly in higher agreement in hypernudges. It might be thought that, on the one hand, individuals tend to take care of interventions by hypernudges; on the other hand, they might agree with this intervention even prohibited one because it is easy for them to stop watching customized advertisements. In contrast, the exact opposite result is shown in No. 7 (education campaign for childhood obesity), which shows a significantly higher disagreement in the hypernudge, even though higher agreement in the original nudge. It seems that individuals would hesitate to present and utilize their children's personal data. No. 16 (security protection) is easier to accept in hypernudges because it is easy to imagine that AI-driven systems are always exposed through viruses and hackers. While several typical reasons can be suggested for these results, it is difficult to give interpretations of why No. 10 (labelling unhealthy food) and 13 (green energy consumption) that are similar to Nos. 1, 2, and 3, have obtained more agreement in hypernudges.

5.2. Effects of intervention categories

The next step has been designed to compare the differences of intervention categories with a one-way analysis of variance. The rates of agreement for each intervention are shown in Table 5.

Table 5. The rates of acceptance for each intervention (%).

Depth	original	hyper
Campaign	73.49	40.27
Man. info.	40.27	68.96
Default	68.96	54.82
Choice arch.	54.82	57.43
Forced	57.43	55.13

Type	original	hyper
Cons.	64.57	46.25
Subcon.	46.25	54.60

Context	original	hyper
Health	56.52	46.46
Ecology	46.46	59.56
Charity	59.56	54.03
Traffic safety	54.03	52.85
Info. security	52.85	50.84

Prosoci.	original	hyper
Social	55.78	51.30
Personal	51.30	61.63

Comparison of five levels of intervention depth only showed marginally significant tendencies. Except for No. 8 (forced), the deeper nudge (Nos. 14 and 15, mandatory choice architecture) is more disagreeable, and the shallowest interventions (Nos. 6 and 7 campaigns) are the most acceptable among them (F=4.991, p=.015). However, this tendency was confirmed in the original nudge, but not in the hypernudge (F=3.117, p=.61). This means that it would be more difficult to consider whether and which interventions by AI-driven hypernudges would be accepted by people because the depth of interventions is not related to the acceptance. Other categories are inconsistent with prior studies and do not show significant differences. Type – original nudge: F=1.534, p=.236; hypernudge: F=1.208, p=.290; context – original nudge: F=.307, p=.867; hypernudge: F=2.133, p=.145; prosociality – original nudge: F=.517, p=.484; hypernudge: F=.121, p=.733. These results show that H3 is partly supported, and only the depth of interventions has significantly different effects in the original nudges.

5.3 The effects of individual difference

Further, we estimated the logistic regression for the five levels of depth of interventions with significant approval rates of the 16 interventions being dependent variables in both the original and the hypernudge. Age, gender (number of male), educational years (of schooling), and political attitude (support for ruling party, opposition, or non-partisan) were used as independent variables.

Notably, Table 6 shows that political attitude has a unique influence on participants' approval of nudges: non-partisan people significantly disapprove three out of five types of original nudges (mandatory information, default, and choice architecture), but significantly approve (negative influence) four types of hypernudges (campaign, mandatory information, default, and forced). Political independents tend to, on the one hand, be more doubtful on general interventions by the original nudges, while on the other hand, being acceptable for new technology and customized data. However, because the rate of non-partisan people was highest among the participants (59.65%), more than half of them had applied for this tendency.

In addition, the age of participants has a negative impact: older people tend to disapprove both one original nudge (default) and three hypernudges (mandatory information, default, and choice architecture). This is in line with Sunstein et al. (2018). Several other factors have significant influence. For example, males tend to less favour original nudges on mandatory information, people who have used a mobile phone for approximately 5 to 10 years tend to disapprove the original nudges on mandatory information, people who have used a mobile phone under one year tend to support the original nudge on choice architecture; also years of education had a significantly negative correlation on the choice architecture. These results suggest that H4 is marginally supported.

Intervention type	Campaigns (Nos. 6 and 7)				Mandatory Information (Nos. 1, 2, 10, and 12)			
	Original		Hypernudge		Original		Hypernudge	
	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)
Constant	2.389*	10.904	-1.209	.299	.978*	2.659	.392	1.479
Age (in years)	.015	1.015 (1.006-1.024)	-.002	.998 (.991-1.006)	.002	1.002 (.969-1.008)	-.015**	.986 (.980-.991)
Gender (male)	-.209	.812 (.615-1.070)	-.028	972 (.762-1.240)	-.208*	.812 (.676-.976)	.021	1.201 (.860-1.213)
Political attitude								
Ruling	-.168	.846 (.609-1.174)	—	—	—	—	—	—
Opposition	.179	1.196 (.799-1.792)	-.167	.846 (.583-1.228)	.209	1.232 (.905-1.677)	.022	1.023 (.775-1.348)
Non-partisan	—	—	.427*	1.532 (1.168-2.010)	-.248*	.781 (.633-.963)	.457**	1.579 (1.295-1.926)
Education (in years)	-.021	.964 (.901-1.032)	-.031	.969 (.905-1.038)	.013	1.013 (.967-1.062)	-.001	.999 (.956-1.034)
Mobile use years								
No use	-.103	.928 (.517-1.665)	.461	1.586 (.968-2.597)	-.235	2.427 (1.053-1.170)	.206	1.229 (.861-1.754)
Under one year	-1.034	.463 (.187-1.149)	.642	1.901 (.911-3.965)	.068	1.070 (.610-1. 778)	.043	1.044 (.640-1.704)
Less than 5 years	-.661	.699 (.434-1.127)	.169	1.184 (.785-1.787)	-.162*	.851 (.611-1.184)	-.051	.950 (.707-1.277)
Less than 10 years	-.340	.695 (.437-1.105)	.287	1.332 (.898-1.978)	-.325*	2.659 (.525-.994)	-.067	.935 (.702-1.246)
Over 10 years	—	—	—	—	—	—	—	—
Obs.	1,192		1,192		2,384		2,384	
Omnibus test	$\chi^2(9)$ =20.211*		$\chi^2(9)$ =21.084*		$\chi^2(9)$ =26.038**		$\chi^2(9)$ =67.401**	

Intervention type	Default (Nos. 3, 4, 5, 9, 11, 13, and 16)				Choice Architecture (Nos.14 and 15)				Forced (No.8)			
	Original		Hypernudge		Original		Hypernudge		Original		Hypernudge	
	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)	β	OR (95% CI)
Constant	.545	1.724	.332	1.394	-2.190*	.112	-.544	.580	-.637	.529	-1.104	.331
Age (in years)	-.004*	.966* (.992-1.000)	-.009**	.991 (.987-.995)	.008	1.008 (1.000-1.0162)	.009*	1.009 (1.001-1.017)	.004	1.004 (.993-1.015)	.000	1.000 (.989-1.011)
Gender (male)	-.052	.950 (.835-1.081)	.060	1.062 (.933-1.209)	.020	1.020 (.791-1.314)	.060	1.062 (.817-1.382)	.027	1.027 (.723-1.457)	-.152	.859 (.611-1.207)
Political attitude												
Ruling	—	—	.218*	1.243 (1.008-1.534)	—	—	.073	1.076 (.792-1.462)	.117	1.124 (.748-1.688)	—	—
Opposition	.088	1.092 (.888-1.343)	—	—	-.049	.952 (.651-1.393)	-.061	.941 (.646-1.370)	.264	1.303 (.775-2.189)	-.367	.693 (.398-1.206)
Non-partisan	-.183*	.833 (.719-.965)	.473**	1.605 (1.324-1.944)	.380*	1.463 (1.103-1.939)	—	—	—		.474*	1.607 (.398-1.206)
Education (in years)	.007	1.007 (.974-1.041)	-.015	.985 (.953-1.018)	-.010	.990 (.923-1.062)	-.068*	.934 (.878-.993)	-.046	.955 (.878-1.038)	.005	1.005 (.912-1.106)
Mobile use years												
No use	-.324*	.724 (.554-.945)	.258	1.294 (.988-1.695)	.116	1.123 (.684-1.845)	.467	1.596 (.925-2.753)	.535	1.707 (.805-3.621)	.503	1.653 (.823-3.321)
Under one year	.037	1.037 (.714-1.506)	-.100	.905 (.625-1.309)	.851*	2.341 (1.051-5.213)	-.378	.685 (.343-1.370)	-.432	.649 (.241-1.748)	.699	2.012 (.735-5.507)
Less than 5 years	-.021	.979 (.783-1.225)	-.035	.966 (.772-1.208)	.263	1.301 (.852-1.989)	.292	1.339 (.860-2.085)	.387	1.473 (.800-2.711)	.197	1.217 (.673-2.200)
Less than 10 years	-.043	.958 (.771-1.190)	-.056	.761 (.484-1.176)	.201	1.223 (.816-1.833)	.409	1.506 (.976-2.323)	.339	1.404 (.775-2.545)	.280	1.323 (.751-2.331)
Over 10 years	—	—	—	—	—	—	—	—	—	—	—	—
Obs.	4,172		4,172		1,192		1,192		596		596	
Omnibus test	$\chi^2(9)$ =28.488**		$\chi^2(9)$ =62.235**		$\chi^2(9)$ =23.179*		$\chi^2(9)$ =29.934**		$\chi^2(9)$ =.9470		$\chi^2(9)$ =17.298*	

* p<.05
** p<.001

Table 6. Estimates of selected individual demographic, political attitude, and mobile usage of selected approval of interventions per five levels of depth: Results of a logistic regression analysis.

6. DISCUSSION AND CONCLUSION

Overall, the results of this study indicated that the approval and disapproval for hypernudges were dramatically different from the original nudges. In this study, it was revealed that hypernudges were not more acceptable than the original nudges. Notably, in hypernudges, while individuals tended to accept the less flexible and forbidden intervention, they rejected the ones that utilized their children's personal data. However, neither typical nor common features were confirmed that could identify the acceptance level of hypernudges, such as categories of interventions, individual sociodemographic factors, political attitudes, and mobile phone usage histories.

As predicted, the deeper the intervention (too much meddling), the less acceptable for people. However, this tendency was seen only in the original nudge. Compared with the original nudge, though only four interventions on 'education campaign for childhood obesity', 'labelling unhealthy food', 'requiring large energy provider to enrol green energy', and 'installing security software' could get more approval in hypernudge, there could be found neither typical nor common features among them. In addition, while several prior studies had investigated that consciousness, contexts, and prosociality of interventions had different effects among people, this study does not recognize the same effects as well. Insignificant effects might stem from a nationality such as Japanese or individual differences that this study did not consider. Sunstein et al. (2018) surveyed the acceptance of nudges in several countries, and it was observed that Japanese had one of the lowest acceptance rates among all. In contrast, Americans, British, and Chinese would favour various types of nudges. Japanese, Hungarians, and Danish tended to hesitate to accept nudges. Although it is esoteric to assert the reasons, we, as Japanese, should take care of the numerous types of interventions, especially by AI-driven artefacts.

This suggests a kind of an alert or dark cloud for the introduction, utilization, and spreading of hypernudges because of neither lower acceptance than the original nudges nor no common and specific traits among the accepted interventions in hypernudges. We should continuously consider the effects of the various categories of hypernudges on various types of people. This study might serve as an onset of prevalence for appropriate AI-driven artefacts.

ACKNOWLEDGEMENTS

This work was supported by the Japan Society for Management Information Grant-in-Aid for SIG 'Monitoring and Control of AI Artefacts' and the Seikei University Grant-in-Aid 2020-21, for the research on 'Monitoring and Control of AI Artefacts: Consideration from Economics, Social, and Legal Perspectives'. I also thank the two anonymous reviewers in ETHICOMP 2020.

REFERENCES

BBS News (2019). Organ donor law change named after Max and Keira, https://www.bbc.com/news/health-47359682

Bovens, L. (2009). The Ethics of Nudge, In T. Yanoff-Grüne and S. O. Hansson. (Eds.) *Preference Change* (pp. 207–209). Dordrecht: Springer.

Bruns, H., Kantorowicz-Reznichenko, E., Klement, K., Jonsson, M. and Rahali, B. (2018). Can Nudges Be Transparent and Yet Effective? *Journal of Economic Psychology*, 65, 41-59.

Diepeveen, S., Ling, T., Suhrcke, M., Roland, M. and Marteau, T.M. (2013). Public acceptability of government intervention to change health-related behaviours: A systematic review and narrative analysis, *BMC Public Health*, 13, 756.

Felsen G, CasteloN, Reiner PB (2013). Decisional Enhancement and Autonomy: Public Attitudes Towards Overt and Covert Nudges. *Judgment and Decision Making*, 8, 202-213.

Furedi, F. (2011). Defending moral autonomy against an army of nudgers, *Spiked.* Retrieved from http://tinyurl.com/6kfafka

Goodwin, T. (2012). Why we should reject 'nudge', *Policy and Politics*, 41(2), 159-182.

Jung, J. Y. and Mellers, B. A. (2016). American Attitudes toward Nudges. *Judgment and Decision Making*, 11, 62-74.

Junghans, A. F., Cheung, T. T. L. and De Ridder, D.T.D. (2015). 'Under consumers' scrutiny: An investigation into consumers' attitudes and concerns about nudging in the realm of health behavior', *BMC Public Health* 15, 336.

Junghans, A. F. and Marchiori, D., and De Ridder, D. T. D. (2016). The Who and How of Nudging: Cross-national Perspectives on Consumer Approval in Eating Behaviour. Unpublished Manuscript, Utrecht: Utrecht University.

IEEE standard association (2018). Affective Computing, *Ethically Aligned Design*, ver. 2nd., pp. 162-181. Retrieved from https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf

Lanzing, M. (2018). Strongly Recommended" Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies, *Philosophical Technology*, 32(3), 549-568.

O'Neil, B. (2011). 'Nick Clegg's sinister nannies are 'nudging' us towards an Orwellian nightmare", *The Telegraph*. http://tgr.ph/g7gLsp.

Reisch, L. A. and Sunstein, C. (2016). Do Europeans Like Nudges? *Judgment and Decision Making*, 1(4), 310-325.

Schnellenbach, J. (2012). Nudges and norms: On the political economy of soft paternalism, *European Journal of Political Economy*, 28, 266-277.

Sunstein, C. R. (2015). *Nudging and Choice Architecture: Ethical Considerations*, (Harvard John M. Olin Discussion Paper Series Discussion Paper No. 809, Jan. 2015, Yale Journal of Regulation.

Sunstein, C. R. (2016). *The Ethics of Influence: Government in the Age of Behavioral Science.* CUP, New York.

Sunstein, C. R. (2018). Misconceptions about nudges, *Journal of Behavioral Economics for Policy*, 2(1), 61-67.

Sunstein, C. R., Reich, L. A., Rauber, J. (2018). A worldwide consensus on nudging? Not quite, but almost, *Regulation & Governance,* 12, 3-22.

Thaler, R. H. and Sunstein C. R. (2008). Nudge: *Improving Decisions about Health, Wealth, and Happiness.* Yale University Press, New Haven, CT.

Yamazaki, Y. (2019). Certified Requirements of Nudging by AI Artefacts: Hypernudges Driven by AI/ML Artifacts will not be Recognized as Nudges, XXVIII AEDEM International Meeting Tokyo (Japan).

Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design, *Information Communication and Society*, 20(1), 118-136.

Weinmann, M., Schneider, C. and Brocke. J. V. (2016). Digital Nudging, *Business and Information Systems Engineering*, 58(6), 433-436.

Wilkinson, T. M. (2012). Nudging and Manipulation, *Political Studies*, 61(2), 341-355.