

---

## A Survey on Investigating Data-Sharing Behavior among STEM Researchers: The Context of a Developing Country

---

WILDA SYAHRI<sup>1</sup>, YUSNAIDAR<sup>1</sup>, MUHAIMIN<sup>2\*</sup>

### Abstract

Based on an extended theory of planned behavior (TPB), this study focused on five exogenous constructs (perceived benefits, facilitating condition, normative belief, behavioral belief, and control belief) that was established to predict behavioral intention and behavior of data-sharing. The research model was tested with 157 Indonesian researchers from science, technology, engineering, and mathematics (STEM). We applied the partial least squares (PLS) techniques to analyze the data. Seven out of eleven hypotheses are supported based on the significance test in through the PLS-SEM procedure. The study's findings show that the proposed extended TPB is a valid and reliable framework for describing the data-sharing behavior of STEM researchers. This paper facilitates an advanced investigation through PLS-SEM approaches toward behavioral intention and behavior regarding data-sharing; the findings are among the first elaborated data-sharing report in the Indonesian context, a country with more than 200 thousand researchers.

### Keywords

STEM researchers, data-sharing, PLS-SEM

---

<sup>1</sup>Faculty of Education and Teacher Training, Universitas Jambi, Jambi, Indonesia

<sup>1</sup>Faculty of Education and Teacher Training, Universitas Jambi, Jambi, Indonesia

<sup>2</sup>Faculty of Pharmacy, Universitas Padjadjaran, Jl. Raya Jatinangor Km 21, Sumedang, Jawa Barat 45363, Indonesia; [muhaimin@unpad.ac.id](mailto:muhaimin@unpad.ac.id)

## Introduction

Data-sharing could be critical in modern science, particularly in disciplines that rely on collaborative processes produced through data-based activities and publications. Data-sharing could be distributed by academicians using a variety of research methodologies, namely experiments, simulations, and surveys (Tenopir et al., 2020). Scientists might have benefited from data-sharing such as findings verification, ideas development, and research extension (Heymann, 2020; Kim & Zhang, 2015; Meyer, 2018). Raw-shared datasets resulting from a project can be used with comparative studies, and more complex analyses can help scientists expand their ideas (Federer et al., 2018; Kim & Zhang, 2015; Zhang et al., 2020). It can also improve collaboration and enable large-scale studies. The research was traditionally published through printed materials, having data-sharing more difficult. However, current researchers could share their publications and dataset easily with the availability of technology such as emails, research applications, publication websites, and data repositories. Although data-sharing is considered beneficial and can make more discoveries in academia, the practice has not been widely adopted and is difficult to understand; therefore, this study was conducted in the context of Indonesia with samples of STEM researchers.

Data-sharing in this study is defined as raw dataset sharing of published articles to understand how researchers decide to release their data. Even though data-sharing is beneficial for all fields of study (e.g., social sciences, business/management, and humanities), this study limits the context in the field of STEM, focusing on determinants that predict Indonesian STEM researchers. The research focuses on five exogenous constructs (perceived benefits, facilitating condition, normative belief, behavioral belief, and control belief) that were developed to predict data-sharing behavioral intention and behavior. Two research objectives were addressed: a) to examine the validity and reliability of the determinants predicting Indonesian STEM researchers' data-sharing behaviors; b) understand and elaborate these determinants predicting Indonesian STEM researchers' data-sharing behaviors.

## Literature Review

### *Potential difficulties in listening comprehension*

Even though data has been an essential forum for scientific communication, data-sharing as a research practice is still limited. (Federer et al., 2018), who gathered data availability statements from 47,593 papers published in a journal by Plos between March 2014 and May 2016, inform that only about 20% of the papers were deposited in a repository in which the publisher policy states the preferred method. Despite its potential to accelerate academic progress, public data-sharing remains relatively uncommon in some disciplines (Houtkoop et al., 2018). Another study (Melero & Navarro-Molina, 2020) that involved researchers from several institutes of the Spanish research council reports some barriers of researchers to open their data, namely lack of a data-sharing culture, fear of being scooped, and confusion between the concepts of the working plan and the data management

plan. In addition, (Zenk-Möltgen et al., 2018), through the theory of planned behavior (TPB), show that data-sharing would improve if journals had explicit data-sharing policies. However, authors also need support from other institutions (their universities, funding councils, and professional associations) to improve data management skills and infrastructures. The results indicate that authors' attitudes, past behavior, social norms, and perceived behavioral control affected their intentions to share data.

Although prior research on data-sharing has yielded valuable results, the improvement in further studies is suggested by broadening the scope of the study to report key affecting determinants, establishing more empirical data. In this study, the TPB was adapted to meet the improvement. The TPB is a well-known social theory that explains how a person's core beliefs impact his/her behavioral intentions and/or behaviors. The idea asserts that an individual's attitudes, subjective norms, and perceived behavioral control impact behavioral intentions and behavior. The TPB, as a motivational theory, has the potential to enlighten how people make decisions based on their personal motives (Ajzen, 1991a). In this study context, this theory refers to a framework to assess data-sharing behaviors. In addition to TPB's exogenous factors (attitudes, subjective norms, and perceived behavioral), this study involved two external variables, namely perceived benefits and facilitating conditions. Similarly, prior researchers have also extended the TPB in their study model (Al-Emran et al., 2020; Kim & Zhang, 2015; Lubida et al., 2015; Nguyen et al., 2019; Sadaf et al., 2012; Yusop et al., 2021). In the model development, the current research relies on self-reported behavior (behavioral intention and behavior) as two expected outcomes; the framework is associated with eleven hypotheses, shown in Figure 1.

### ***Perceived benefits***

Perceived benefits regarding data-sharing could promote better communication, planning, and future demand in research (Powell et al., 2021). In this study, we define the perceived benefits as the value that academics obtain from presenting working performance, getting a higher rate of citations, and improving images and credits from data-sharing activities. Researchers' opinions of data-sharing advantages would improve their positive attitudes since they value credits, performance, and reputation (Kim & Zhang, 2015; Yoon & Kim, 2020). Studies using data that had been placed in repositories received a better quantity of citations than studies with data that had not been deposited; thus, data-sharing promotes academic incentives and higher citations (Park & Wolfram, 2017; Piwowar & Vision, 2013). When researchers have more citations, they might realize that the activity generates the quality of their work (Rowhani-Farid et al., 2017; Sanderson et al., 2017). However, (Thelwall & Kousha, 2017) informed that an insignificant relationship appears between data-sharing and research impact or citation. Therefore, perceived benefits are suggested to support more positive attitudes toward data-sharing. To understand the phenomena, two hypotheses are established based on the prediction of impacts resulting from perceived benefits toward attitudes and behavioral intention:

- H1: Perceived benefits significantly affect attitudes toward data-sharing.
- H2: Perceived benefits positively predict behavioral intention regarding data-sharing.

### *Facilitating conditions*

The extent to which people feel that accessible technological and organizational support systems will drive them to utilize a system referred to as the facilitating conditions (Venkatesh et al., 2003). The availability of infrastructure is one of the facilitating conditions indicated in the research on open data-sharing (Zuiderwijk et al., 2020). In the context of this study, facilitating conditions relate to the availability of essential resources, namely data repositories for researchers' data-sharing, a website for data storage, emails synchronization, and research applications (Harper & Kim, 2018; Mooney & Newton, 2012). Prior studies in data-sharing have investigated facilitating conditions to significantly influence attitudes toward data-sharing and behavioral intention (Harper & Kim, 2018; Mooney & Newton, 2012; Zuiderwijk et al., 2020). These prior studies informed that facilitating conditions encouraged STEM researchers' attitudes and behavioral intention regarding data-sharing. Two hypotheses were proposed to disclose the role of facilitating conditions towards attitudes and behavioral intention.

H3. Facilitating conditions significantly affect attitudes.

H4: Facilitating conditions are a significant predictor of behavioral intention.

### *Subjective norms*

One of the fundamental constructs in TPB is subjective norms, described as a construct that measures the extent to which an individual feels that people who are close and important to them want them to behave in a certain way (Ajzen, 1991). Subjective norms in this study refer to the influence of other researchers that are connected to the participants of the study (Kim & Zhang, 2015); perceptions that other researchers feel that the respondents should do data-sharing to share their publication(s). Previous studies found that subjective norms significantly predicted attitudes and behavioral intention for knowledge sharing (Alajmi, 2012; Khalil et al., 2014; Mousa et al., 2019). Regarding data-sharing, prior studies highlighted the important role of subjective norms in affecting attitudes and behavioral intention regarding data-sharing (Kim & Nah, 2018; Kim & Stanton, 2016; Kim & Zhang, 2015). Therefore, two hypotheses were proposed regarding subjective norms on data-sharing among Indonesian STEM researchers.

H5. Subjective norms positively affect attitudes.

H6. Subjective norms are significantly correlated with behavioral intention.

### *Attitudes*

Attitudes refer to an individual's opinions of a specific behavior (Ajzen, 1991b). In this study, attitudes are described as the level to which STEM researchers who responded to the survey have negative or positive opinions on data-sharing behavior (Howe et al., 2018; Kim & Zhang, 2015). Researchers who published their articles in journals would prefer to open their data when they have positive opinions on the sharing (Joo & Kim, 2017; Zhu, 2020). On the other hand, if they have negative perceptions toward data-sharing, the

---

researchers will not open their data (Kim & Nah, 2018; Kim & Stanton, 2016; Kim & Zhang, 2015). Kim and Zhang (Kim & Zhang, 2015) reported that attitudes on data-sharing that included several sub-constructs (perceived career benefit, career risk, and effort) had significant relationships toward behavioral intention to open or share research data. In this research, the correlation between attitudes and behavioral intention regarding data-sharing was assessed. One hypothesis was proposed:

H7. Attitudes toward data-sharing significantly affect behavioral intention.

### ***Perceived behavioral control***

Ajzen (Ajzen, 1991b) introduced the construct ‘perceived behavioral control’ into TPB as a determinant of behavioral intention and actual behavior. On a conceptual basis, perceived behavioral control refers to the person’s belief that the behavior in question is under his or her control—but operationally, perceived behavioral control is often assessed by the ease or difficulty of the behavior (Ajzen, 1991b). In this study context, the ease of data-sharing was highlighted in the perceived behavioral control aspect that was expected to significantly influence attitudes, behavioral intention, and behavior regarding data-sharing among STEM researchers in Indonesia. Prior studies explored the relationship by informing perceived behavioral control as a strong predictor of attitudes, behavioral intention, and behavior regarding data or knowledge sharing (Hau & Kang, 2016; Houtkoop et al., 2018; Lubida et al., 2015).

H8. Perceived behavioral control positively affects attitudes.

H9. Perceived behavioral control is significantly correlated with behavioral intention

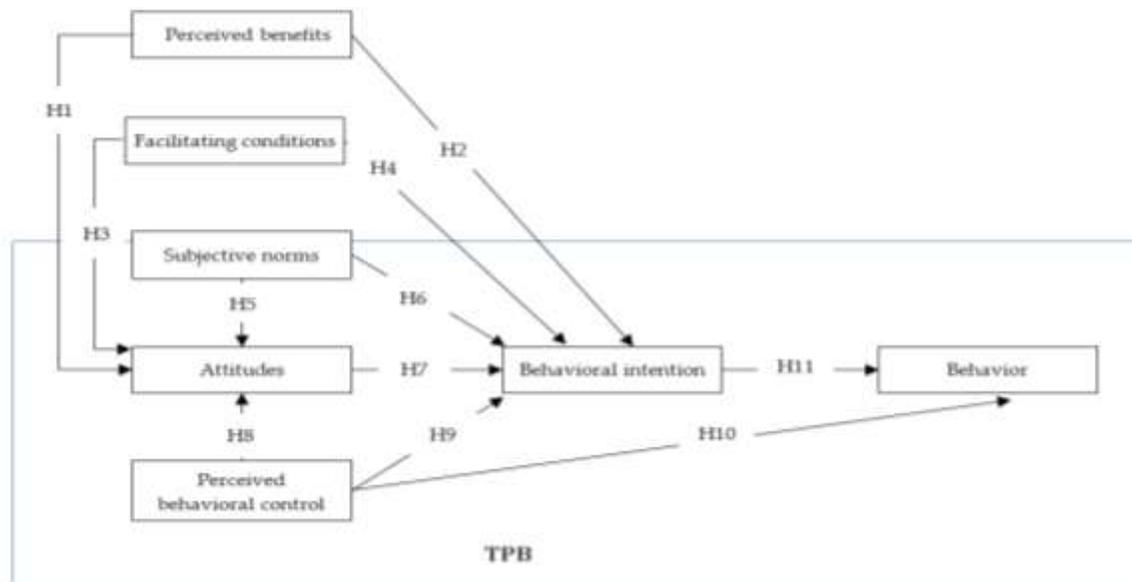
H10. Mediated by behavioral intention, perceived behavioral control has an indirect effect on behavior.

### ***Behavioral intention and behavior***

STEM researchers that engage in data-sharing behavior make research data available through data repositories, public online spaces, supplemental materials, or personal communication channels upon request. Individual motivational variables, namely attitudes, subjective norms, and perceived behavioral control, are thought to be captured by behavioral intention. At the same time, the decision to engage in the behavior is a direct determinant of the actual conduct (Ajzen, 1991a). In this study, behavioral intention refers to the motivational factors affecting STEM researchers to open their data when they publish their articles in journals. The stronger the data-sharing intention among researchers, the more likely they open their data. Prior studies have explored this relationship and found that behavioral intention was a significant predictor of behavior on data-sharing (Jeon et al., 2011; Kim & Nah, 2018; Kim & Stanton, 2016). Therefore, the data-sharing behavior could be predicted by behavioral intention in the current study:

H11. Behavioral intention is a strong predictor of behavior.

Figure 1. The proposed model to examine determinants affecting data-sharing behavior, STEM researchers



## Methodology

This study adapted TPB in which we extended the theory by adding two external variables: perceived benefits and facilitating conditions to support the original TPB variables; attitudes, subjective norms, perceived behavioral control, behavioral intention, and behavior regarding data-sharing. Eleven hypotheses were included in this study (Fig 1). The current study could facilitate beneficial understanding for all stakeholders like librarians and researchers to support scientific data-sharing in STEM. Empirically, we applied a survey approach to validate the research model and to investigate the constructs and connections. It also yielded broad conclusions regarding academics' data-sharing habits across fields.

## Population and sample

STEM researchers in Indonesia are the target population of the current study. We obtained data from Indonesian science and technology index for its sampling pool. Based on the data, 223,662 listed scholars in 4 main major fields of disciplines classed by the Indonesian science and technology index scholar database. The four major discipline categories included science and engineering, health, social, and art and humanity. We categorized science and engineering, and health fields as the STEM researchers ( $n= 73,673$ ) as the population in this study; 50,820 researchers from science and engineering major and 22,853 from health. The target population was represented in the sampling pool for this study. Thus, the sampling was extrapolated to the entire population. A probability random sampling approach was used to choose the survey respondents (Creswell, 2013). From the Indonesian science and technology scholar database, respondents were randomly selected

from researchers who work at three institutions. In addition, they should have published at least one paper within the previous two years. The minimal sample size for this research is 107 participants based on \*G power assessment, sample size calculation for a correlation between two continuous variables with five predictors and two outcome variables (Erdfelder et al., 2009).

### *Instrumentations*

Twenty-seven indicators were adapted from previous studies (Kim & Nah, 2018; Kim & Stanton, 2016; Kim & Zhang, 2015). In the development initiation, a panel of 5 experts and 5 users reviewed the indicators in three focus group discussions as part of the content validity process. Based on the discussions, we revised a few items as a refinement effort to make the indicators suitable for the study samples; three indicators were dropped. Indicators with Likert scales of 1 to 5 (strongly disagree to strongly agree) were applied for five constructs' indicators (perceived benefits, facilitating conditions, attitudes, subjective norms, and perceived behavioral control). Meanwhile, different measures (Likert scales from 1–5, never to always) were implemented for behavioral intention and behavior. We piloted the instrument to thirty-seven STEM researchers to assess its reliability through the Cronbach's alpha test in SPSS 23.0; as a result, all indicators were reliable with no alpha values less than .700

### *Data collection and analysis*

The instrument was distributed to three Indonesian institutions from 28 April 2021 to 15 August 2021. From more than 1000 researchers listed in STEM disciplines (science and engineering, and health), we obtained only 157 responses or around 15% return rate. The assumption for the low responses is that because most STEM invited researchers did not share their data. In the first part of the questionnaire, we provided the option of whether they have ever shared their data or not; if not, the questionnaire cannot be filled in. Ninety-three respondents are females, while 64 are males. Thirty-one respondents have doctoral degrees; 126 respondents are master's degree holders. In addition, 124 respondents are from science and engineering and 33 from health.

This study used a structural equation modeling (SEM) method to explore connections among components in the research model. Because the goal of this study is exploratory rather than confirmatory, a component-based SEM employing partial least squares (PLS) was selected (Hair, et al., 2019). SmartPLS 3.3.3 was utilized as the data analysis tool in this study. The PLS analysis was conducted in two phases. A measurement model was applied to see how closely the indicators for each latent variable referred to a similar conceptual construct (Hair, et al., 2019). A structural model was addressed to evaluate the connections between the constructs (Dijkstra & Henseler, 2015; Leguina, 2015).

## Findings

### *Measurement model*

For the measurement model, we first examined reflective indicators loadings. This initial phase refers to a data computation in PLS-SEM through the PLS algorithm, informing that the values higher than 0.708 are the recommended values (Habibi et al., 2020; Hair, et al., 2019). However, loading values above 0.500 could still be retained (Noor et al., 2019; Ogbeibu et al., 2021). The threshold refers to the constructs elaborated 50% of the variance or above, facilitating the items' reliability. Using SmartPLS 3.3.3, the data were statistically processed to report the loading of all items. No items dropped since all loading values are above 0.500. Table 1 exhibits sufficient values of the loadings from all twenty-four items. The lowest loading was gained from behavior (B4; 0.7080), while the highest value was derived from behavioral intention (BI2; 0.9290). Afterwards, reliability and internal consistency were assessed; the assessments were facilitated by calculating Cronbach alpha, rho\_A, and Composite Reliability (CR). In Table 1, values (+0.700) of Cronbach's alpha, rho\_A, and CR confirm all constructs' reliability and internal consistency (Ogbeibu et al., 2021).

**Table 1.** *Reflective indicator loadings, internal consistency reliability, and convergent validity*

Construct	Item	Load	$\beta$	rho_A	CR	AVE	VIF
Attitudes	AT1	0.7960	0.7430	0.7810	0.8290	0.5500	1.467
	AT2	0.8310					1.646
	AT3	0.6890					1.982
	AT4	0.6330					1.821
Behavior	B1	0.7980	0.8420	0.8520	0.8870	0.6120	1.836
	B2	0.7810					1.659
	B3	0.8100					1.847
	B4	0.7080					1.707
	B5	0.8080					2.027
Behavioral intention	BI1	0.9080	0.8150	0.8240	0.9150	0.8430	1.898
	BI2	0.9290					1.898
Facilitating conditions	FC1	0.8540	0.8120	0.8310	0.8880	0.7250	1.725
	FC2	0.7950					1.736
	FC3	0.9020					2.276
Perceived benefits	PB1	0.8770	0.8410	0.8400	0.9050	0.7600	2.259
	PB2	0.8420					1.979
	PB3	0.7990					1.821
	PB4	0.8110					1.938
Perceived behavioral control	PBC1	0.9020	0.8530	0.8660	0.9000	0.6940	2.729
	PBC2	0.8880					2.557
	PBC3	0.8240					1.593
Subjective norms	SN1	0.9130	0.8150	0.8670	0.8890	0.7280	2.138
	SN2	0.7690					1.575
	SN3	0.8720					1.977

The average variance extracted (AVE) values of greater than 0.500 have an indication of satisfactory convergent validity (Habibi et al., 2021; Henseler, 2017); All AVE data show good values of more than 0.500, ranging from 0.5500 to 0.8430 (Table 1). To avoid the multicollinearity issue, VIF values were examined. VIF values informed in Table 1 indicate that the study's data are free of multicollinearity issues since they all are below 4 (Hair et al., 2010). In addition, the discriminant validity was addressed by examining the heterotrait-monotrait ratio (HTMT), suggested as the highest boundary criterion for discriminant validity (Hair, et al., 2019). Consistent with (Hair, et al., 2019) notes, the HTMT results of Table 2 indicate that all constructs are different, supporting the model's discriminant validity.

**Table 2.** *Heterotrait-monotrait ratio (htmt) of all constructs, below 0.900*

	Attitudes	Behavior	Behavioral intention	Facilitating conditions	Perceived behavioral control	Perceived benefits	Subjective norms
Attitudes	0.7420						
Behavior	0.6970	0.7820					
Behavioral intention	0.4010	0.5140	0.9180				
Facilitating conditions	0.5670	0.5390	0.4640	0.8520			
Perceived behavioral control	0.4790	0.5240	0.6220	0.3490	0.8720		
Perceived benefits	0.5410	0.5070	0.5500	0.4290	0.6700	0.8330	
Subjective norms	0.5610	0.5330	0.3250	0.4310	0.4460	0.3430	0.8530

### ***Structural model***

The current study applied consistent bootstrapping in SmartPLS 3.3.3 with 5,000 subsamples in estimating the structural model. Prior to the elaboration of the structural model, we estimated the model's fit criteria by providing the elaboration of Standardized Root Mean Square Residual (SRMR), the squared Euclidean Distance (d\_ULS), and the Geodesic Distance (dG) that are based on prior extant research (Dijkstra & Henseler, 2015; Hu & Bentler, 1998, 1999; Ogbeibu et al., 2021). To achieve the model's fits, the SRMR value should be less than 0.800. As exhibited in Table 3, the SRMR value (0.060) refers to an appropriate model fit for the current study and indicates that the proposed model has an adequate value of model fit. The d\_ULS and dG have also performed good values of 2.002 and 0.698, respectively. To examine the structural model, studies (Ringle et al., 2020; Sarstedt et al., 2016) recommend assessing measures such as statistical significance,  $f^2$  (effect sizes),  $R^2$  (coefficient of determination), and  $Q^2$  (predictive relevance).

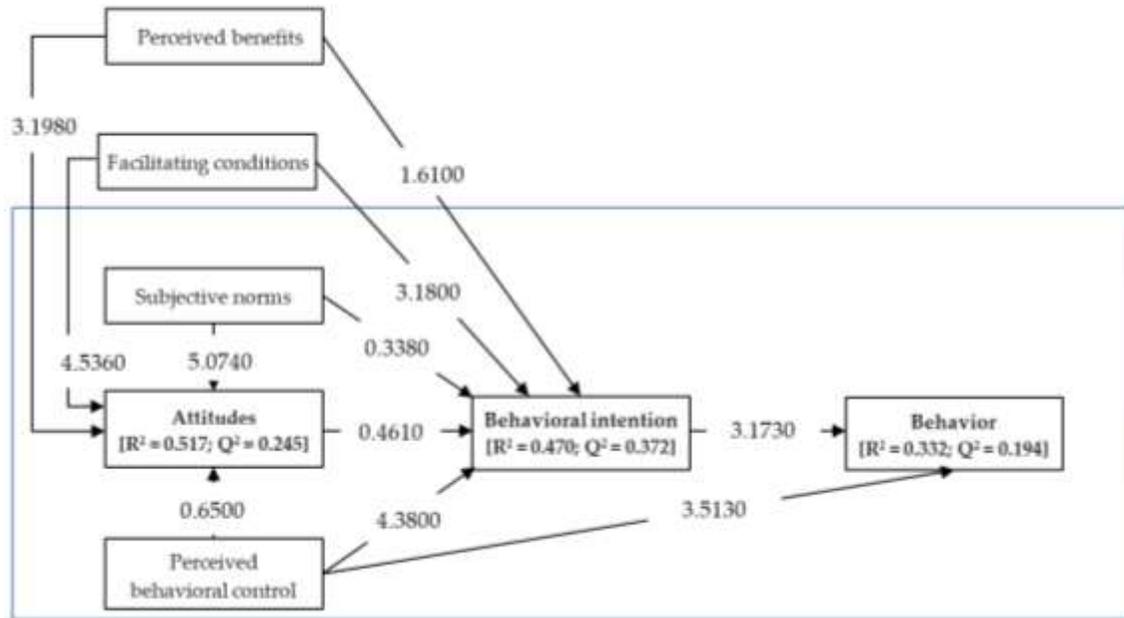
For the statistical significance, the results inform that seven out of eleven relationships are positively significant. The seven supporting hypotheses are H1, H3, H4,

H5, H9, H10, H11. The strongest relationship between subjective norms and attitudes emerges with a t value of 5.0740 ( $\beta = 5.0740$ ;  $p < 0.001$ ), supporting H5. On the other hand, the weakest relationship that confirms H11 is between behavioral intention and behavior with the value of 3.1730 ( $\beta = 0.3080$ ;  $p < 0.05$ ). Four relationships are not statistically significant that are perceived benefits  $\rightarrow$  behavioral intention, subjective norms  $\rightarrow$  behavioral intention, attitudes  $\rightarrow$  behavioral intention, and perceived behavioral control  $\rightarrow$  attitudes. In terms of effect sizes ( $f^2$ ), Ringle et al. (2018) advocate that  $f^2$  scores of 0.02, 0.15, and 0.35 suggest small, medium, and large effects; no large effects are found in the current study. The effect sizes of all exogenous constructs to all endogenous constructs are informed in detail in Table 3. The coefficient of determination  $R^2$  is defined as the value measuring the predictive accuracy—the computation functions as the correlation of squares among endogenous variables. Figure 2 exhibits the endogenous constructs'  $R^2$ , values of 0.67, 0.33, and 0.19 refer to strong, moderate, and weak predictive relevance. A  $Q^2$  value of more than 0 for endogenous constructs indicates good predictive accuracy. Recommended by prior research (Hair, et al., 2019; Ogbeibu et al., 2021) for reporting models' predictive accuracy,  $Q^2$  results of attitudes (0.245), behavioral intention (0.372), and behavior (0.194) indicate a good predictive relevance for all exogenous variables.

**Table 3.** *Statistical significances, effect sizes, and model's fits*

H	Path	$\beta$	p	Sig.	$f^2$	Remarks	Model's fits
H1	Perceived benefits $\rightarrow$ Attitudes	0.2690	0.0010	Yes	0.076	Small	SRMR 0.060
H2	Perceived benefits $\rightarrow$ Behavioral intention	0.1660	0.1070	No	0.024	Small	d_Uls 2.002
H3	Facilitating conditions $\rightarrow$ Attitudes	0.2960	0.0000	Yes	0.131	Medium	d_G 0.698
H4	Facilitating conditions $\rightarrow$ Behavioral intention	0.2710	0.0010	Yes	0.089	Small	
H5	Subjective norms $\rightarrow$ Attitudes	0.3170	0.0000	Yes	0.149	Medium	
H6	Subjective norms $\rightarrow$ Behavioral intention	-0.0250	0.7350	No	0.001	No effect	
H7	Attitudes $\rightarrow$ Behavioral intention	-0.0420	0.6450	No	0.002	No effect	
H8	Perceived behavioral control $\rightarrow$ Attitudes	0.0540	0.5160	No	0.003	No effect	
H9	Perceived behavioral control $\rightarrow$ Behavioral intention	0.4480	0.0000	Yes	0.188	Medium	
H10	Perceived behavioral control $\rightarrow$ Behavior	0.3330	0.0000	Yes	0.102	Medium	
H11	Behavioral intention $\rightarrow$ Behavior	0.3080	0.0020	Yes	0.087	Small	

**Figure 2.** Bootstrapping ( $t$  value and  $R^2$ ) and blindfolding results ( $Q^2$ ) of the structural model to examine determinants affecting data-sharing behavior, stem researchers



## Discussion

The scale development was done to produce a valid and reliable instrument (Davis et al., 2018). Firstly, previous literature sources were analyzed for the instrumentation process, generating and initiating twenty-seven indicators. The instrument was then validated through face and content validity by discussing all indicators with experts and users for contextual and setting suitability. Face validity and content validity, though qualitative methods are important steps in survey validation to evaluate to what extent the instrument can measure the study purposes (Wynd, et al., 2003). From these processes, three items were dropped. Afterwards, the study's instrument was piloted and analyzed through Cronbach's alpha test in SPSS 23. Twenty-four indicators were included for the main data collection. The data, 157 responses, were computed for the measurement model within the PLS-SEM approaches (J. F. Hair et al., 2020). No indicators were eliminated in the measurement model processes. SRMR,  $d_U$ s, and  $d_G$  were reported to understand the model's fits before the structural model assessment was conducted with the remaining indicators (24 items). Validation of a survey instrument is an important activity in the research process (Ferketich et al., 1993).

This valid and reliable instrument was intended to investigate how the extended TPB factors established in this study affect STEM researchers' attitudes, behavioral intention, and behavior regarding data-sharing. Three constructs are significant in predicting attitudes toward data-sharing: perceived benefits, facilitating conditions, and subjective. When STEM researchers perceive more benefits, more supporting facilitating conditions such as accessible technological and organizational support systems for data-sharing, and

more supports from their important friends and colleague, they are more likely to have positive attitudes toward data-sharing. The significance of facilitating conditions in predicting attitudes toward data-sharing supports prior studies (Harper & Kim, 2018; Mooney & Newton, 2012; Zuiderwijk et al., 2020). In addition, perceived benefits were also reported as factors that predict attitudes in research (Kim & Zhang, 2015; Park & Wolfram, 2017; Piwowar & Vision, 2013; Powell et al., 2021). Prior studies in knowledge sharing also informed that subjective norms are a key predictor for attitudes (Alajmi, 2012; Khalil et al., 2014; Mousa et al., 2019). However, perceived behavioral control does not significantly impact attitudes that contradict prior studies (Hau & Kang, 2016; Houtkoop et al., 2018; Lubida et al., 2015).

Regarding behavioral intention, facilitating conditions that refer to the facilitation of important resources, such as data repositories, website storage, and other facilities for data-sharing, significantly predict behavioral intention. The findings agree with prior studies (Harper & Kim, 2018; Mooney & Newton, 2012; Zuiderwijk et al., 2020), encouraging that facilitating conditions support STEM researchers' behavioral intention regarding data-sharing. Perceived behavioral control also significantly predicts behavioral intention; this significance confirms studies investigating the relationships between the two (Hau & Kang, 2016; Houtkoop et al., 2018; Lubida et al., 2015). However, two core factors of TPB (perceived behavioral control and attitudes) and one extended construct (perceived benefits) are insignificant in predicting behavioral intention. Unlike the previous findings that reveal the significances of perceived behavioral control (Hau & Kang, 2016; Houtkoop et al., 2018; Lubida et al., 2015), attitudes (Kim & Nah, 2018; Kim & Stanton, 2016; Kim & Zhang, 2015), and perceived benefits (Kim & Zhang, 2015; Rowhani-Farid et al., 2017; Sanderson et al., 2017), the insignificances reported in this study should be more comprehended with bigger samples to see the reality, fact, and reasons on why the insignificances emerged.

The last endogenous construct, behavior, was hypothesized to be predicted by perceived behavioral control and behavioral intention. Based on the analyses with the structural model, perceived behavioral control is significantly related to behavior that is similar to previous research (Hau & Kang, 2016; Houtkoop et al., 2018; Lubida et al., 2015). The more respondents believe that they can handle the data-sharing process, the better they do the data-sharing activities. In addition, behavioral intention is also reported to significantly predict behavior. Individual STEM researchers' decision in engaging within the behavior significantly determines the behavior of data-sharing. The motivational factors triggering STEM researchers to do data-sharing can improve the behavior (Jeon et al., 2011; Kim & Nah, 2018; Kim & Stanton, 2016; Kim & Zhang, 2015). Data-sharing behavior was significantly determined by the researchers' intention.

### **Conclusion**

The findings of the study include several implications. Based on TPB, this research considers attitudes, behavioral intention, and behavior as outcome variables. The data resulting from this study quantifies data-sharing behavior. The findings inform the measurement of the three endogenous variables that can work as important outcome variables. The findings need for future studies to elaborate more on the data-sharing

behavior. Several practical implications can be considered for all stakeholders and policymakers providing data-storage services. The study's findings consider that STEM researchers perceived behavioral control and behavioral intention are important to bridge researchers' data-sharing behaviors. Related stakeholders should have a good decision on how to ease researchers for appropriate data-sharing tools and management, providing researchers better facilities. Reports revealed that data-sharing activities need a lot of hard work and effort (Kim & Nah, 2018; Kim & Zhang, 2015). To encourage data-sharing, researchers should be supported to manage their data sets, thus permitting them to share their data with their colleagues more easily.

The small sample obtained by the current study is one of the main limitations. We have tried to send the survey questionnaire to more than 1000 STEM researchers; however, the 15 % responses might indicate that STEM researchers are not accustomed to data-sharing (Kim & Stanton, 2012). Other limitations refer to the specific field of study, STEM, that general fields of study with bigger samples can be addressed for future research. Demographic information involvement (gender, institution, and fields of study) in the data analysis through tests of differences like t-test, analysis of variance, multi-analysis of variance, and multi-group analysis can also be recommended for future studies. The method is also limited to quantitative data, specifically surveys; therefore, both quantitative and qualitative approaches are recommended. For instance, an interview can be addressed to gain a more in-depth understanding of the data-sharing phenomena. Experimental studies can be more beneficial to see how a treatment works in the data-sharing activities. The model can also be extended for future researchers with more outcomes and predicting variables.

#### ***Disclosure statement***

No potential conflict of interest was reported by the authors.

#### ***Acknowledgments***

The research/publication of this article was funded by Universitas Jambi.

#### **References**

- Ajzen, I. (1991a). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I. (1991b). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Alajmi, B. M. (2012). The intention to share: Psychological investigation of knowledge sharing behaviour in online communities. *Journal of Information and Knowledge Management*, 11(3). <https://doi.org/10.1142/S0219649212500220>
- Al-Emran, M., Arpaci, I., & Salloum, S. A. (2020). An empirical examination of continuous intention to use m-learning: An integrated model. *Education and Information Technologies*, 25(4), 2899–2918. <https://doi.org/10.1007/s10639-019-10094-2>
- Creswell, J. (2013). Qualitative, quantitative, and mixed methods approaches. In *Research design*.

- Davis, D., Chen, G., Hauff, C., & Houben, G. J. (2018). Activating learning at scale: A review of innovations in online learning strategies. *Computers and Education*, 125. <https://doi.org/10.1016/j.compedu.2018.05.019>
- Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis*, 81. <https://doi.org/10.1016/j.csda.2014.07.008>
- Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Federer, L. M., Belter, C. W., Joubert, D. J., Livinski, A., Lu, Y. L., Snyder, L. N., & Thompson, H. (2018). Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLoS ONE*, 13(5). <https://doi.org/10.1371/journal.pone.0194768>
- Ferketich, S., Phillips, L., & Verran, J. (1993). Development and administration of a survey instrument for cross-cultural research. *Research in Nursing & Health*, 16(3). <https://doi.org/10.1002/nur.4770160310>
- Habibi, A., Yaakob, M. F. M., Mukminin, A., Muhaimin, M., Prasajo, L. D., Yusop, F. D., & Muzakkir, M. (2021). Teachers' digital technology access to motivation, skills and use: a structural equation modeling study. *Aslib Journal of Information Management, ahead-of-p*(ahead-of-print). <https://doi.org/10.1108/ajim-11-2020-0382>
- Habibi, A., Yusop, F. D., & Razak, R. A. (2020). The role of TPACK in affecting pre-service language teachers' ICT integration during teaching practices: Indonesian context. *Education and Information Technologies*, 25(3), 1929–1949. <https://doi.org/10.1007/s10639-019-10040-2>
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). Multivariate Data Analysis: A Global Perspective. In *Multivariate Data Analysis: A Global Perspective* (Vol. 7th).
- Hair, J. F., Astrachan, C. B., Moisescu, O. I., Radomir, L., Sarstedt, M., Vaithilingam, S., & Ringle, C. M. (2020). Executing and interpreting applications of PLS-SEM: Updates for family business researchers. *Journal of Family Business Strategy*. <https://doi.org/10.1016/j.jfbs.2020.100392>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. In *European Business Review* (Vol. 31, Issue 1, pp. 2–24). <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair, J. F., Sarstedt, M., & Ringle, C. M. (2019). Rethinking some of the rethinking of partial least squares. *European Journal of Marketing*, 53(4), 566–584. <https://doi.org/10.1108/EJM-10-2018-0665>
- Harper, L. M., & Kim, Y. (2018). Attitudinal, normative, and resource factors affecting psychologists' intentions to adopt an open data badge: An empirical analysis. *International Journal of Information Management*, 41. <https://doi.org/10.1016/j.ijinfomgt.2018.03.001>
- Hau, Y. S., & Kang, M. (2016). Extending lead user theory to users' innovation-related knowledge sharing in the online user community: The mediating roles of social capital and perceived behavioral control. *International Journal of Information Management*, 36(4). <https://doi.org/10.1016/j.ijinfomgt.2016.02.008>

- Henseler, J. (2017). *Partial Least Squares Path Modeling* (pp. 361–381).  
[https://doi.org/10.1007/978-3-319-53469-5\\_12](https://doi.org/10.1007/978-3-319-53469-5_12)
- Heymann, D. L. (2020). Data sharing and outbreaks: best practice exemplified. In *The Lancet* (Vol. 395, Issue 10223). [https://doi.org/10.1016/S0140-6736\(20\)30184-7](https://doi.org/10.1016/S0140-6736(20)30184-7)
- Houtkoop, B. L., Chambers, C., Macleod, M., Bishop, D. V. M., Nichols, T. E., & Wagenmakers, E. J. (2018). Data Sharing in Psychology: A Survey on Barriers and Preconditions. *Advances in Methods and Practices in Psychological Science*, 1(1).
- Howe, N., Giles, E., Newbury-Birch, D., & McColl, E. (2018). Systematic review of participants' attitudes towards data sharing: A thematic synthesis. In *Journal of Health Services Research and Policy* (Vol. 23, Issue 2).
- Hu, L. T., & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification. *Psychological Methods*, 3(4).
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1).
- Jeon, S. H., Kim, Y. G., & Koh, J. (2011). Individual, social, and organizational contexts for active knowledge sharing in communities of practice. *Expert Systems with Applications*, 38(10). <https://doi.org/10.1016/j.eswa.2011.04.023>
- Joo, Y. K., & Kim, Y. (2017). Engineering researchers' data reuse behaviours: A structural equation modelling approach. *Electronic Library*, 35(6).  
<https://doi.org/10.1108/EL-08-2016-0163>
- Khalil, T., Atieh, K., Mohammad, A. U., & Bagdadlian, E. F. (2014). Examining the Social and Technical Factors Influencing School Teachers Knowledge Sharing Intentions in a Teachers Online Professional Community. *The Electronic Journal of Knowledge Management*, 12(3).
- Kim, Y., & Nah, S. (2018). Internet researchers' data sharing behaviors: An integration of data reuse experience, attitudinal beliefs, social norms, and resource factors. *Online Information Review*, 42(1). <https://doi.org/10.1108/OIR-10-2016-0313>
- Kim, Y., & Stanton, J. (2012). Institutional and Individual Influences on Scientists' Data Sharing Practices. *The Journal of Computational Science Education*, 3(1).
- Kim, Y., & Stanton, J. M. (2016). Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, 67(4). <https://doi.org/10.1002/asi.23424>
- Kim, Y., & Zhang, P. (2015). Understanding data sharing behaviors of STEM researchers: The roles of attitudes, norms, and data repositories. *Library and Information Science Research*, 37(3), 189–200. <https://doi.org/10.1016/j.lisr.2015.04.006>
- Leguina, A. (2015). A primer on partial least squares structural equation modeling (PLS-SEM). *International Journal of Research & Method in Education*, 38(2), 220–221.
- Lubida, A., Pilesjö, P., Espling, M., & Runnström, M. (2015). Applying the theory of planned behavior to explain geospatial data sharing for urban planning and management: Cases from urban centers in Tanzania. *African Geographical Review*, 34(2).  
<https://doi.org/10.1080/19376812.2014.892433>
- Md Noor, S., Rasoolimanesh, S. M., Jaafar, M., & Barghi, R. (2019). Inscription of a destination as a world heritage site and residents' perceptions. *Asia Pacific Journal of Tourism Research*, 24(1). <https://doi.org/10.1080/10941665.2018.1541183>

- Melero, R., & Navarro-Molina, C. (2020). Researchers' attitudes and perceptions towards data sharing and data reuse in the field of food science and technology. *Learned Publishing*, 33(2). <https://doi.org/10.1002/leap.1287>
- Meyer, M. N. (2018). Practical Tips for Ethical Data Sharing. *Advances in Methods and Practices in Psychological Science*, 1(1). <https://doi.org/10.1177/2515245917747656>
- Mooney, H., & Newton, M. P. (2012). The Anatomy of a Data Citation: Discovery, Reuse, and Credit. *Journal of Librarianship and Scholarly Communication*, 1(1).
- Mousa, T. S., Jameel, A. S., & Ahmad, A. R. (2019). The impact of attitude, subjective norm and information communications technology on knowledge sharing among academic staff. *International Journal of Psychosocial Rehabilitation*, 23(2). <https://doi.org/10.37200/IJPR/V23I2/PR190325>
- Nguyen, T. M., Nham, P. T., & Hoang, V. N. (2019). The theory of planned behavior and knowledge sharing: A systematic review and meta-analytic structural equation modelling. *VINE Journal of Information and Knowledge Management Systems*, 49(1). <https://doi.org/10.1108/VJIKMS-10-2018-0086>
- Ogbeibu, S., Jabbour, C. J. C., Gaskin, J., Senadjki, A., & Hughes, M. (2021). Leveraging STARA competencies and green creativity to boost green organisational innovative evidence: A praxis for sustainable development. *Business Strategy and the Environment*.
- Park, H., & Wolfram, D. (2017). An examination of research data sharing and re-use: implications for data citation practice. *Scientometrics*, 111(1).
- Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 2013(1). <https://doi.org/10.7717/peerj.175>
- Powell, K. R., Deroche, C. B., & Alexander, G. L. (2021). Health Data Sharing in US Nursing Homes: A Mixed Methods Study. *Journal of the American Medical Directors Association*, 22(5). <https://doi.org/10.1016/j.jamda.2020.02.009>
- Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2020). Partial least squares structural equation modeling in HRM research. *International Journal of Human Resource Management*, 31(12), 1617–1643. <https://doi.org/10.1080/09585192.2017.1416655>
- Rowhani-Farid, A., Allen, M., & Barnett, A. G. (2017). What incentives increase data sharing in health and medical research? A systematic review. *Research Integrity and Peer Review*, 2(1).
- Sadaf, A., Newby, T. J., & Ertmer, P. A. (2012). Exploring pre-service teachers' beliefs about using Web 2.0 technologies in K-12 classroom. *Computers and Education*, 59(3), 937–945. <https://doi.org/10.1016/j.compedu.2012.04.001>
- Sanderson, S. C., Brothers, K. B., Mercaldo, N. D., Clayton, E. W., Antommara, A. H. M., Aufox, S. A., Brilliant, M. H., Campos, D., Carrell, D. S., Connolly, J., Conway, P., Fullerton, S. M., Garrison, N. A., Horowitz, C. R., Jarvik, G. P., Kaufman, D., Kitchner, T. E., Li, R., Ludman, E. J., ... Holm, I. A. (2017). Public Attitudes toward Consent and Data Sharing in Biobank Research: A Large Multi-site Experimental Survey in the US. *American Journal of Human Genetics*, 100(3).
- Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, 69(10), 3998–4010. <https://doi.org/10.1016/j.jbusres.2016.06.007>

- Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLoS ONE*, *15*(3).
- Thelwall, M., & Kousha, K. (2017). Do journal data sharing mandates work? Life sciences evidence from Dryad. *Aslib Journal of Information Management*, *69*(1).
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*, *27*(3), 425–478. <https://doi.org/10.2307/30036540>
- Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. In *Western Journal of Nursing Research* (Vol. 25, Issue 5).
- Yoon, A., & Kim, Y. (2020). The role of data-reuse experience in biological scientists' data sharing: an empirical analysis. *Electronic Library*, *38*(1).
- Yusop, F. D., Habibi, A., & Razak, R. A. (2021). Factors affecting Indonesian preservice teachers' use of ICT during teaching practices through theory of planned behavior. *SAGE Open*, *11*(2), 215824402110275. <https://doi.org/10.1177/21582440211027572>
- Zenk-Möltgen, W., Akdeniz, E., Katsanidou, A., Naßhoven, V., & Balaban, E. (2018). Factors influencing the data sharing behavior of researchers in sociology and political science. *Journal of Documentation*, *74*(5). <https://doi.org/10.1108/JD-09-2017-0126>
- Zhang, Z., Song, S., Yu, J., Zhao, W., Xiao, J., & Bao, Y. (2020). The Elements of Data Sharing. *Genomics, Proteomics and Bioinformatics*, *18*(1).
- Zhu, Y. (2020). Open-access policy and data-sharing practice in UK academia. *Journal of Information Science*, *46*(1). <https://doi.org/10.1177/0165551518823174>
- Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLoS ONE*, *15*(9 September). <https://doi.org/10.1371/journal.pone.0239283>
- 

### Biographical Notes

**Dr. WILDA SYAHRI** is a senior lecturer at the Faculty of Education and Teacher Training, Universitas Jambi, Indonesia.

**Dr. YUSNAIDAR** is a senior lecturer at the Faculty of Education and Teacher Training, Universitas Jambi, Indonesia.

**Dr. MUHAIMIN** is a full professor at Faculty of Pharmacy, Universitas Padjadjaran, Jl. Raya Jatinangor Km 21, Sumedang, Jawa Barat 45363, Indonesia; [muhaimin@unpad.ac.id](mailto:muhaimin@unpad.ac.id)