Journal of Physics D: Applied Physics

TOPICAL REVIEW • OPEN ACCESS

Raw material ‘criticality’—sense or nonsense?

To cite this article: M Frenzel *et al* 2017 *J. Phys. D: Appl. Phys.* **50** 123002

OPEN ACCESS
IOP Publishing

Journal of Physics D: Applied Physics

J. Phys. D: Appl. Phys. 50 (2017) 123002 (18pp)

doi:10.1088/1361-6463/aa5b64

Topical Review

Raw material 'criticality'—sense or nonsense?

M Frenzel[1,4], J Kullik[1,2], M A Reuter[1] and J Gutzmer[1,3]

[1] Helmholtz-Zentrum Dresden-Rossendorf, Institute Freiberg for Resource Technology, Chemnitzer Str. 40, 09599 Freiberg, Germany
[2] Institute of Political Science, Professorship of International Relations, Technische Universität Chemnitz, Thüringer Weg 9, 09126 Chemnitz, Germany
[3] Institute for Mineralogy, Technische Universität Bergakademie Freiberg, Brennhausgasse 14, 09599 Freiberg, Germany

E-mail: m.frenzel@hzdr.de

Received 26 October 2016, revised 20 January 2017
Accepted for publication 23 January 2017
Published 20 February 2017

Abstract

The past decade has seen a resurgence of interest in the supply security of mineral raw materials. A key to the current debate is the concept of 'criticality'. The present article reviews the criticality concept, as well as the methodologies used in its assessment, including a critical evaluation of their validity in view of classical risk theory. Furthermore, it discusses a number of risks present in global raw materials markets that are not captured by most criticality assessments. Proposed measures for the alleviation of these risks are also presented.

We find that current assessments of raw material criticality are fundamentally flawed in several ways. This is mostly due to a lack of adherence to risk theory, and highly limits their applicability. Many of the raw materials generally identified as critical are probably not critical. Still, the flaws of current assessments do not mean that the general issue of supply security can simply be ignored. Rather, it implies that new assessments are required. While the basic theoretical framework for such assessments is outlined in this review, detailed method development will require a major collaborative effort between different disciplines along the raw materials value chain.

In the opinion of the authors, the greatest longer-term challenge in the raw materials sector is to stop, or counteract the effects of, the escalation of unit energy costs of production. This issue is particularly pressing due to its close link with the renewable energy transition, requiring more metal and mineral raw materials per unit energy produced. The solution to this problem will require coordinated policy action, as well as the collaboration of scientists from many different fields—with physics, as well as the materials and earth sciences in the lead.

Keywords: supply security, high-tech metals, critical metals, critical mineral resources, strategic raw materials

(Some figures may appear in colour only in the online journal)

[4] Author to whom any correspondence should be addressed.

1. Introduction

In recent years, concerns about supply security have led to a resurgence of interest in mineral and metal availability, particularly in highly import-dependent industrialised countries such as Germany, Japan, and the United States. The rapidly increasing number of studies published in this field clearly reflects this trend (e.g. Glöser *et al* 2015, Graedel and Reck 2015, Helbig *et al* 2016, NSTC 2016). After a period of relatively low raw-material prices following the end of the Cold War (Humphreys 1995), the rise of China as the World's major producer and consumer of minerals, metals and energy has led to a tightening of global markets, and concomitant fears of inadequate supply (Buijs and Sievers 2011a, 2011b).

Ultimately, the trigger for the first new wave of government studies into non-fuel raw materials supply (NRC 2007, EU Commission 2010) was the reduction of Chinese export quotas for rare earth elements (REEs) by ~25 % from 2007 to 2009 (Tse 2011). China's use in 2010 of export restrictions for REEs as a political instrument against Japan (over tensions in the East China Sea) further aggravated these concerns (Buijs and Sievers 2011a, 2011b). At the time, China accounted for more than 95% of global primary REE output (Hedrick 2010, Cordier 2011). Therefore, the effects of these quota reductions were widely felt across the globe, mostly in terms of significant price increases.

The concept of raw material criticality is at the core of the recent debate. It is the aim of this review to introduce this concept to a physical sciences audience, while critically discussing both its limitations and the shortcomings of current assessments in some detail. In addition, we will briefly consider those complexities of the raw material markets not captured by current assessments of criticality. We argue that generalised criticality assessments can be misleading, and might consequently not be useful to highlight the greatest (economic) risks inherent in current markets. In contrast to the simplified material-by-material approach used in current assessments, we try to identify over-arching issues and discuss a number of possible approaches for the alleviation of the corresponding risks.

2. Raw material criticality

Despite its central role in the current debate, it is still difficult to give a precise definition for raw material criticality. This is mainly due to two factors: (1) the variability in definitions and assessment methodologies used and (2) the vagueness of the definitions given in many studies (see Erdmann and Graedel 2011, Glöser *et al* 2015). For instance, Graedel and Nuss (2014) define criticality as 'the quality, state or degree of being of the highest importance', while according to Gleich *et al* (2013), 'criticality denotes the extent of current and future risks associated with a certain metal'. Raw materials with a high criticality rating are called *critical*.

Irrespective of this variability in definitions, two aspects are widely regarded as crucial for the identification of critical raw materials: (1) the vulnerability of a given consumer to supply disruptions, i.e. the importance of the raw material under consideration, and the consequent impact of supply shortfalls, and (2) the likelihood for the occurrence of such disruptions, often called 'supply risk' (NRC 2007, Erdmann and Graedel 2011, Graedel and Reck 2015). Critical raw materials must have a high score on both dimensions when compared to other raw materials (e.g. EU Commission 2010).

Glöser *et al* (2015) clearly demonstrated that the basic 2D structure of most assessments reflects a fundamental equivalence between criticality and the classic definition of risk (e.g. Cox 2009). Other studies also hint at this equivalence (Buijs *et al* 2012, Graedel and Nuss 2014, Helbig *et al* 2016). Given this observation, it is possible to define the criticality of a raw material more precisely as *a measure of the (economic) risk arising from its utilisation (incl. production, use, and end-of-life) for a specific consumer over a certain period*. A consumer can be anything from a single company or technology, to a national or multi-national economy (Graedel *et al* 2012).

Because this definition ties criticality to risk theory, and thus allows for a rigorous treatment within the framework of quantitative risk analysis (Cox 2009), we use it as the basis for the present review. Readers should note that *risk*, both in the definition given above as well as in quantitative risk analysis, is not limited to being economic in nature (see Cox 2009). It may also be societal or environmental, as discussed in more detail in section 3 below. Therefore, our definition preserves the 'holistic' intentions of the criticality concept (e.g. Helbig *et al* 2016). Nevertheless, we note that most criticality assessments focus exclusively on *economic* aspects in their practical implementation, and this is why we specifically included them in our definition.

As indicated above, and stressed by most authors (see Erdmann and Graedel 2011, Graedel and Reck 2015), the criticality of a given raw material is strongly dependent on the consumer and time-period under consideration. Criticality is therefore by no means an inherent property of a material itself, although the similarity of the term to the criticality concept in the physical sciences, where it generally refers to systems near a transformation of state, might at first suggest such a relationship (see Bradshaw *et al* 2013).

It is also important to note that virtually all criticality assessments are limited to *non-fuel* raw materials, even though their structure would in principle allow for the inclusion of fuels (coal, gas, oil). In this review, the term 'raw material(s)' therefore explicitly refers to *non-fuel* raw materials.

2.1. Assessment method

The most common procedure for criticality assessments is to compile sets of different indicators into aggregate scores for both the supply risk and vulnerability dimensions, and then plot them against each other to delimit the field of critical raw materials. Figure 1 illustrates such a plot. This approach is an abstraction from the classical risk matrix (see Glöser *et al* 2015), and was originally introduced by the National Research Council (NRC 2007). As mentioned above, some studies merely use a single final dimension (e.g. BGS 2015, NSTC 2016), while others introduce a third one to reflect e.g. environmental issues (Graedel *et al* 2012).

The detailed assessment process differs widely across different studies. In particular, there are large differences in the specific choice of indicators, their respective weightings in the final aggregate scores (Erdmann and Graedel 2011, Graedel and Reck 2015), and the procedures used for the aggregation of these scores (Erdmann and Graedel 2011, Glöser *et al* 2015).

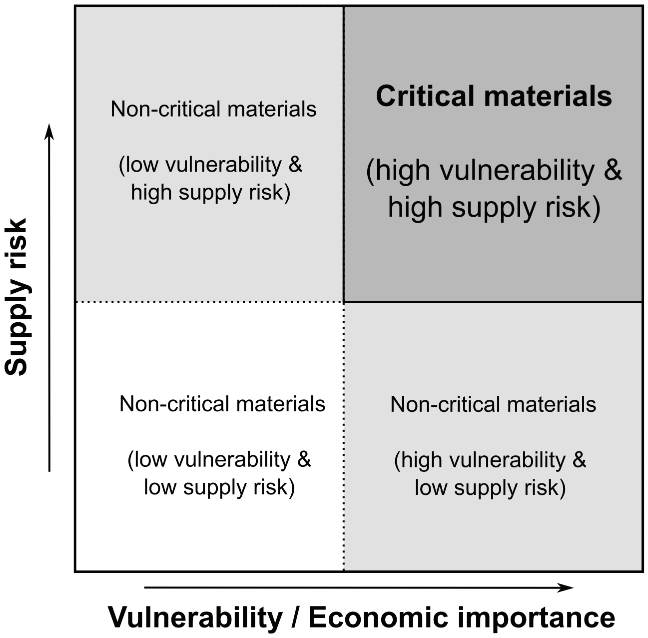

Figure 1. Schematic criticality plot showing the way in which critical raw materials are identified in many studies (e.g. EU Commission 2014).

While some of the differences in the specific choices of indicators between studies certainly reflect differences in the systems under consideration (i.e. the specific consumer, see above), there is still much variation between studies with a similar scope. In the following two subsections, we give a brief overview of the different indicators used for the vulnerability and supply risk dimensions in recent studies, as well as the ways in which they are aggregated. To ensure comparability, we focus specifically on assessments dealing with consumption at the level of national or multi-national economies. A more detailed discussion would be beyond the scope of this article. We nevertheless note that the general principles also apply to assessments at other organisational levels (e.g. single companies; Lapko *et al* 2016).

A critical discussion of the suitability of current assessments to adequately capture the economic risks associated to raw material supply and consumption follows in section 3.

2.2. Indicators for vulnerability

We mentioned earlier that the vulnerability dimension (sometimes called 'economic importance') is intended to provide a measure of the (economic) impact of supply restrictions on the consumer of interest. Economic impact in this sense is perhaps best defined as the additional monetary cost arising from a demand/supply imbalance (Helbig *et al* 2016). As discussed in more detail in the next section, there are many potential contributors to this cost, and due to the great complexity of the global manufacturing and trade system, it is inherently difficult to measure. This is why criticality studies generally use one or more easier-to-measure indicators to assign an estimated vulnerability value. The general assumption behind all such indicators is that they are either directly or inversely related to economic impact. However, concrete choices differ greatly between different studies, as illustrated by the overview in table 1. Readers will note that certain vulnerability indicators are also used for the supply risk dimension in some studies. In these cases, they are not used as vulnerability indicators.

Table 1 summarises the vulnerability indicators used in nine recent assessments (see Helbig *et al* 2016). It shows that substitutability is the most widely used vulnerability indicator. This is a good example of an indicator that is inversely related to vulnerability: the more easily a material can be substituted, the lower the impact of supply restrictions should be, since substitutes are generally less likely to be affected by restrictions at the same time as the raw material of interest. However, substitutability alone is insufficient to assess economic impact. This is, because it does not provide a measure of the magnitude of the costs associated to supply restrictions.

The second most frequently used types of vulnerability indicators are those referring to the value of the products affected, the value of the utilised materials, or the spread of utilisation (table 1). These are good examples of indicators directly related to economic impact. This is particularly clear for the first two, i.e. the value of affected products and the value of utilised materials. Both can in principle be expected to be proportional to the potential economic impact of supply shortfalls. As such, they complement substitutability.

For many of the other indicators listed in table 1, e.g. company concentration of production or consumption volume, the relationship to economic impact is not as clear-cut. This is particularly true for those indicators also used to assess *supply risk* in certain studies. This issue is discussed in more detail in section 3 below.

2.3. Indicators for supply risk

The supply risk dimension is intended to provide a measure for the probability of the occurrence of supply restrictions. Again, this is difficult to quantify due to the complexities of global supply-chains and the many possible events that could result in a tightening of supply relative to demand—from mine-closures and natural disasters to political and military conflicts, or the introduction of new consumer technologies. It is not surprising, therefore, that criticality studies again resort to a set of indicators assumed to be related directly or inversely to supply risk. Just as for vulnerability, particular choices vary greatly between individual studies. This is illustrated by the overview given in table 2 for 15 recent assessments (data from Achzet and Helbig 2013).

Clearly, the most widely used supply risk indicator is country (or company) concentration of production, usually measured using the Herfindahl–Hirschman-Index (HHI; Hirschman 1964). This is often weighted in some way by 'country risk', e.g. using the Worldwide Governance Indicator (WGI; World Bank 2015). The assumption behind the use of these indicators is that the more concentrated the production of a given raw material is in a small number of countries, specifically those with poor governance, the higher will be the probability of supply disruptions due to political or military conflicts in, or with, these countries (e.g. Graedel *et al* 2012). A similar logic applies to company concentration in mining

Table 1. Indicators used for vulnerability/economic importance for national-level studies.

Indicator category	Relation with vulnerability[a]	Frequency of use	Means of measurement/units
Substitutability[b]	Inverse	6	Qualitative
Value of products affected	Direct	3	For example weighted value of mega-sectors in which raw material is used; USD, EUR, % of GDP
Future demand/supply ratio[b]	Direct	3	Qualitative
Value of the utilised materials	Direct	2	USD, USD/kg, % of GDP
Spread of utilisation	Direct	2	% of population utilising material
Change in demand share[b]	Dep. on meas.	2	% p.a.
Import dependence[b]	Direct	2	%
Target groups demand share	Direct	1	%
Strategic importance	Direct	1	Qualitative
Ability to innovate	Inverse	1	Qualitative
Change in imports	Dep. on meas.	1	% p.a.
Company concentration of production[b]	Direct	1	Qualitative
Consumption volume	Direct	1	Kg
Mine production change	Dep. on meas.	1	% p.a.
Recyclability[b]	Inverse	1	Qualitative

[a] Direct relationship: indicator value high for high vulnerability; Inverse relationship: indicator value high for low vulnerability; Dep. on meas. = either direct or inverse depending on measurement.
[b] Indicator also used for supply risk.
Note: data from table 3 of Helbig *et al* (2016) for nine recent assessments.

Table 2. Supply risk indicators (national level).

Indicator	Relation with supply risk[a]	Frequency of use	Means of measurement/units
Country concentration of production	Direct	12	Herfindahl–Hirschman-Index
Country governance	Dep. on def.	10	Qualitative, index
Depletion time	Inverse	9	Years
By-product dependency	Direct	7	%
Company concentration in mining corporations[b]	Direct	5	Herfindahl–Hirschman-Index
Demand growth[b]	Direct	5	Qualitative, ratio
Import dependence[b]	Direct	3	%, net value
Recycling/recycling potential[b]	Inverse	3	Tons
Substitutability[b]	Inverse	3	Qualitative
Volatility of commodity prices	Direct	2	USD/kg, EUR/kg
Exploration degree	Inverse	1	USD, EUR
Production costs in extraction	Direct	1	USD, EUR
Stock keeping	Inverse	1	%
Market balance	Direct	1	Tons
Mine/refinery capacity	Inverse	1	%
Future market capacity	Inverse	1	%
Investment in mining	Inverse	1	USD/t, EUR/t
Climate change vulnerability	Direct	1	Qualitative
Temporary scarcity	Direct	1	Qualitative
Risk of strategic use	Direct	1	Qualitative
Abundance in earth's crust	Direct	1	Ppm

[a] Direct relationship: indicator value high for high vulnerability; Inverse relationship: indicator value high for low vulnerability; Dep. on meas. = either direct or inverse depending on definition.
[b] Also used as a vulnerability indicator.
Note: data from table 3 of Achzet and Helbig (2013) for 15 recent assessments.

corporations as a supply risk indicator. Both are good examples of *direct* indicators.

In contrast, depletion time is a good example of an indicator with an assumed *inverse* relationship to supply risk: The greater the expected time-frame for the depletion of current reserves, it is argued, the lower the probability of shortages within a given period (e.g. Graedel *et al* 2012).

Of the other indicators, by-product dependency requires particular explanation. Many elements considered essential for modern technologies are produced as by-products from the extraction of other metals (Campbell 1985), their so-called hosts. A good example is indium, which is mostly recovered from sulphidic zinc ores (Schwarz-Schampera 2014, Werner *et al* 2017). Many authors argue that this dependence on the production of the host metals results in a higher probability of supply disruptions (e.g. Nassar *et al* 2015). This is because the supply of the by-products is in principle limited by the amount of the host metal extracted every year, and as such, might not be able to adjust to (rapid) increases in demand, resulting in shortages. It is argued that the fraction of a given raw material produced as a by-product can be used as an indicator for the probability of supply disruptions (Nassar *et al* 2015).

Last but not least, the demand growth and import dependence indicators deserve mention. The general argument with respect to demand growth is that a higher value should be accompanied by a higher probability for the development of market-imbalances and consequently a higher supply risk. Similarly, greater import dependence should lead to a greater probability of a country's supply being affected by political or military conflicts elsewhere in the world, also leading to higher supply risk.

Due to their relatively rare usage, the remaining indicators listed in table 2 do not warrant detailed discussion.

2.4. Indicator aggregation and calculation of criticality scores

Different procedures are used to aggregate indicators into the final vulnerability and supply risk scores. This sub-section gives a brief overview of the most important aggregation methods used in different studies. It also introduces the different procedures used to derive an overall criticality score in some studies.

Indicator aggregation methods may conveniently be subdivided into the following three groups:

(1) *Additive*—scores are the sum of (weighted and/or scaled) indicator values (e.g. Graedel *et al* 2012, Panousi *et al* 2016).
(2) *Multiplicative*—scores are the product of (weighted and/or scaled) indicator values (e.g. vulnerability score by EU Commission 2014).
(3) *Mixed*—scores are mixed sums and products of (weighted and/or scaled) indicator values (e.g. the WGI-HHI indicator used by EU Commission 2014).

The numerical results of these aggregation methods will clearly differ widely, even if the same indicator values are used.

Having arrived at the final scores, some studies use simple thresholding to delimit a field of critical raw materials in a so called *criticality plot* (e.g. EU Commission 2014; see figure 1), while others combine either the two dimensions, or the initial indicators, into an overall criticality score that allows for a sorting of materials from highest to lowest criticality. Again, the calculation of this overall score is done differently in different studies. The following methods are in use:

(1) *Sum*—the overall criticality score is the (weighted) sum of the two dimensions or individual indicators (e.g. Morley and Eatherly 2013, Moss *et al* 2013, BGS 2015).
(2) *Euclidean distance*—the overall criticality score is the Euclidean distance from the origin of the criticality plot (sometimes including an additional dimension; e.g. Graedel *et al* 2012).
(3) *Geometric mean*—the overall criticality score is the geometric mean of the two dimensions or of individual indicators (e.g. NSTC 2016).

We note that this list might not be exhaustive. Some studies do not provide sufficient details of the methods they use to aggregate overall scores.

3. Limitations and shortcomings of current approaches

In this section, we discuss the major limitations and shortcomings of current approaches to criticality assessments. We start by examining the relationship of criticality to classical risk and decision theory, and in particular, to quantitative risk analysis (Cox 2009). This gives us a theoretical framework within which an objective evaluation of different aspects of current assessments is possible. Several authors have previously voiced criticism concerning the validity of current assessments (e.g. Buijs and Sievers 2011a, 2011b, Graedel and Reck 2015, Reuter *et al* 2015, Reuter 2016). However, these previous contributions mostly focussed on specific aspects of the assessments, such as the specific choice of indicators, rather than developing their criticism from a more general theoretical framework. The main inspiration for this section was a recent article by Glöser *et al* (2015). Much of the following two subsections represents a summary of their work.

3.1. Relationship to classic risk and decision theory

Classic risk theory stems from Bayesian decision theory (BDT), as used in engineering problems involving probabilistic uncertainties about the *state of nature* (Benjamin and Cornel 1970, Cox 2009). Benjamin and Cornell (1970) define state of nature as 'some factor which is not known with certainty, but on which the consequences of a decision depend', e.g. whether or not there exists a potentially active geological fault in the bedrock below the proposed site for a nuclear power plant. The general principle of BDT is as follows: sets of possible actions, $\{a_i\}$, and states of nature, $\{s_j\}$, are defined in the beginning. To every action-state of nature-pair, a numeric value describing its outcome, usually expressed in monetary terms as a cost (negative) or profit (positive), is assigned using a so-called utility function, $u(a_i, s_j)$. Assuming that the probabilities for the possible states of nature, $\{p(s_j)\}$, are known, expectation values for the utility of each action can be calculated as:

$$E(u|a_i) = \sum_j p(s_j) \times u(a_i, s_j). \tag{1}$$

The action a_i yielding the highest expectation value is the most preferable (Benjamin and Cornell 1970).

While the above is a description of the simplified case where the amount of information on outcomes and probabilities is fixed, it is nevertheless sufficient to arrive at the classical definition of risk. In particular, the risk, R, associated to an uncertain event (accident, hazard), e_i, may in analogy to (1) be expressed as the expectation value of the associated cost, A, i.e.:

$$R(e_i) = p(e_i) \times A(e_i) \tag{2}$$

where $p(e_i)$ is the probability for the occurrence of event e_i. The major difference of the cost function, $A(e_i)$, to the utility function, $u(a_i, s_j)$, described above is that the sign is reversed, such that greater risks are identified by a greater positive value even though their expected impact is negative. It should also be noted that the cost function, $A(e_i)$, will generally represent the sum over a number of different consequences of event e_i. Furthermore, the cost function is not necessarily limited to monetary costs. For instance, when public health risks are considered, a useful measure of cost may be quality-adjusted life-years (Cox 2009). For environmental risks, yet other measures are appropriate (Suter 2006). There is a long-standing debate whether or not these different measures can be harmonised (e.g. Ackerman and Heinzerling 2002). In fact, much of the current discussion about counter-measures to climate change centres around the correct pricing of carbon dioxide emissions (Grubb and Newbery 2007). While a detailed discussion of these issues is beyond the scope of this article, we note that all the different risks (economic/environmental/societal) included in the criticality definition of section 2 are treatable within the general framework of classic risk theory. However, the harmonisation of different cost functions might not be possible such that separate assessments for e.g. economic and environmental risks might be necessary. Although we focus on economic risks in the following discussion, the same principles also apply to the other types of risk.

Another important note on (2) is that it implicitly refers to a particular time frame. This is because the value of $p(e_i)$ generally depends on the length of time under consideration (Cox 2009). A simple example illustrates this assertion: say the event under consideration is the failure of a moving part in a machine, e.g. an axle. Then $p(e_i)$ should clearly be greater for greater time-periods since the probability of failure increases with increasing wear.

In introductory texts, (2) is often re-written as (e.g. Cox 2009):

$$\text{Risk} = \text{likelihood} \times \text{consequence} \tag{3}$$

or some semantically equivalent expression. As Glöser *et al* (2015) demonstrated, the criticality concept as first introduced by the National Research Council (NRC 2007) is clearly an abstraction from this definition of risk. In particular, criticality is implicitly constructed to be directly proportional to the risk associated with the consumption of a given raw material.

Glöser *et al* (2015) also showed that the two dimensions, supply risk and vulnerability commonly used to identify critical raw materials can be matched directly to likelihood and consequence, respectively, as defined by (2) and (3). Provided we assume that the respective relationships of the supply risk and vulnerability indicators with likelihood and consequence are linear we have:

$$\begin{aligned} \text{Criticality} &= \text{supply risk} \times \text{vulnerability} \\ &= (S \times \text{likelihood}) \times (V \times \text{consequence}) = K \times \text{Risk} \end{aligned} \tag{4}$$

where S, V and K are proportionality constants, with $K = S \times \mathrm{V}$ (see Glöser *et al* 2015). The supply risk dimension does therefore clearly *not* constitute a measure of risk in the classical sense, but merely of probability or likelihood. As such, the term is highly misleading. We therefore replace it by *disruption probability* in the following.

Before moving on to the detailed discussion of recent studies, it is necessary to stress that (2)–(4) are greatly simplified, since they refer to single types of events with a fixed effect size only.

To arrive at a generalised algebraic form for raw material criticality, one needs to consider what kinds of events actually contribute to the risk criticality describes. Unfortunately, the discussion of this fundamental aspect is exceedingly vague in most studies. Nevertheless, there is a general consensus that, irrespective of the underlying causes, the main events of interest are price-hikes (type 1), and severe physical disruptions of supply (type 2), potentially causing the shut-down of production of dependent products (e.g. NRC 2007, Buijs and Sievers 2011a, 2011b).

Readers will note that: (a) the effect size of type 1 events is continuous (price can increase by an arbitrary factor, e.g. 2, 3, 4 etc), with a corresponding probability distribution, and b) type 2 events will generally coincide with (albeit prohibitive) increases in price for the consumer. As such, type 2 events may simply be taken to form the upper end of the scale of effect sizes (prohibitively high cost, equivalent to no availability). This means both types of events may in principle be captured on a single continuous scale of effect sizes. A generalised algebraic form for the criticality, C, of a given raw material can then be written as the expectation value of the vulnerability:

$$C = \int_0^\infty p(x) \cdot v(x)\, \mathrm{d}x \tag{5}$$

where x is the effect size (e.g. the factor by which price increases), and $p(x)$ and $v(x)$ are the corresponding disruption probability (as a probability density) and vulnerability, respectively. Note that vulnerability may also be *negative* in cases where price decreases/supply increases (i.e. where $x < 1$). If vulnerability is expressed as a monetary cost, then C would be in units of e.g. U.S. dollars (USD), since x is dimensionless, as is $p(x)$.

However, (5) still neglects the time for which an event lasts and therefore implies a comparison between events of identical duration. If this is not true, then (5) must be extended to include event duration, t, in addition to effect size:

$$C = \int_{t=0}^{t=t_0} \int_{x=0}^{x=\infty} p(x, t) \cdot v(x, t)\, \mathrm{d}x\, \mathrm{d}t \tag{6}$$

where t_0 denotes the total time period under consideration (e.g. 5 years). Because $p(x,t)$ is in units of $1/t$, the resultant C in (6) is still expressed in monetary units. Note that Glöser *et al* (2015) only derived (4), but did not cite any generalised expressions equivalent to (5) or (6). These are first shown here.

If a criticality assessment were to focus on societal or environmental rather than economic risks, a similar expression would result for C consisting of the integral of the product of disruption probability times vulnerability (in appropriate units) over effect size and time. Only the final units of C will be different if they are not harmonised to monetary costs, e.g. quality-adjusted life-years in the case of public health risks.

Given these quantitative definitions, we are now in a position to evaluate some of the formal aspects of criticality assessments, as well as their underlying logic. In the following two sub-sections, we focus first on the identification and ordering of critical raw materials (section 3.2), and second, on the effects of the addition or omission of assessment dimensions (section 3.3). We then briefly discuss indicator choice and aggregation (section 3.4) and give a final summary in section 3.5. We focus mostly on formal aspects since a detailed discussion of individual indicators and their aggregation is beyond the scope of this article. As shown below, the focus on formal issues is entirely sufficient for our purposes.

3.2. Identification and ordering of critical raw materials

The first important question when evaluating any criticality assessment must be whether its general structure honours the formal requirements set by classic risk theory outlined above. In particular, if an overall criticality score is computed, does it comply with (2) to (6) such that the resultant ordering of raw materials reflects their true criticality? A related question is whether criticality plots (see figure 1) are a suitable tool for the identification of critical raw materials. Since the correct ordering and identification of risk factors is a key requirement for the effective allocation of resources to reduce risk, the following discussion will mostly focus on the effects of different methods on overall criticality order.

It is clear from (5) that an overall criticality score should be calculated as the integral over the product of vulnerability and disruption probability[5] as a function of effect size, x. *Not a single criticality study* in the available literature complies with this requirement. Instead, the dependence of both event probability and vulnerability on x is ignored, and only a single value compiled for each dimension. This is a gross over-simplification. In particular, it implies the comparison of only a single type of event with given x, which is itself never specified; or independence of $p(x)$ and $v(x)$ from x (see section 3.1).

[5] That is, unless non-linear relationships exist between the computed scores and the true vulnerability and disruption probability. In this case, a more complicated aggregation procedure might be required. However, this important detail is not addressed in any recent study, and we will therefore ignore it for the purpose of the following discussion. The majority of workers do not appear to be aware of the close relationship between risk and criticality, and what it entails (see Glöser *et al* 2015), not to speak of these more intricate aspects.

This gross simplification alone highly limits the applicability of current criticality studies, and a justification for it is *never given*.

Even if we assume that the simplification in (4) is legitimate, it is interesting to ask whether current studies at least comply with this simpler form, i.e. whether they calculate overall criticality scores as the product of vulnerability and disruption probability. On a criticality plot (disruption probability against vulnerability) contours of constant criticality should then have the shape of hyperbolas (see Glöser *et al* 2015), as illustrated schematically by the boundary of the criticality field in figure 2(a).

Referring back to section 2.4, it appears that there is *not a single study* where the simple product is used, and a justification of this evident failure to honour (4) is also *never given*. We note, however, that the geometric mean (NSTC 2016) produces exactly the same criticality order and contour shapes as the simple product, although absolute criticality values differ. Therefore, its use does not have a negative impact on qualitative results (ordering). But what effect do the other methods have? In particular, how do their results differ from those of the correct method?

In order to answer these questions, it is helpful to consider a graphical illustration of how the sum and Euclidean distance methods affect the shape of criticality contours in a criticality plot. This is done in a schematic manner in figures 2(b) and (c). The boundaries between the critical and non-critical fields in all plots are designed to go through the centre of the plot as a reference point. The figure clearly illustrates how the criticality field is extended towards the axes as first the sum method (2(b)) and then the Euclidean distance method (2(c)) are used to calculate criticality scores. This means that raw materials plotting close to the axes, and therefore having low true criticality, become more likely to be identified as critical materials—an undesirable result.

In fact, if there is a negative correlation between disruption probability and vulnerability the resultant criticality order will be significantly different from the true criticality order. The broken line in figure 2 illustrates this point. The filled circle at the centre of the line marks the material with the highest true criticality. Moving away from this point towards the axes, true criticality decreases.

It is easily seen that the sum method would result in equal criticality scores for all materials on the line, i.e. a random ordering, while the Euclidean distance method would actually result in a reversal of the true criticality order. The latter result would be considered '*worse-than-useless*' sensu Cox (2008, 2009).

Since perfect negative correlation represents a limiting scenario, we might expect the sum method to always do better than a random assignment of scores for real data. The Euclidean distance method, on the other hand, would be expected to do worse-than-random in an appreciable number of cases. The overall results of any study using the Euclidean distance method (e.g. Graedel *et al* 2012) should therefore be regarded with the greatest scepticism. It is clearly the worst possible choice of method for the calculation of overall criticality scores.

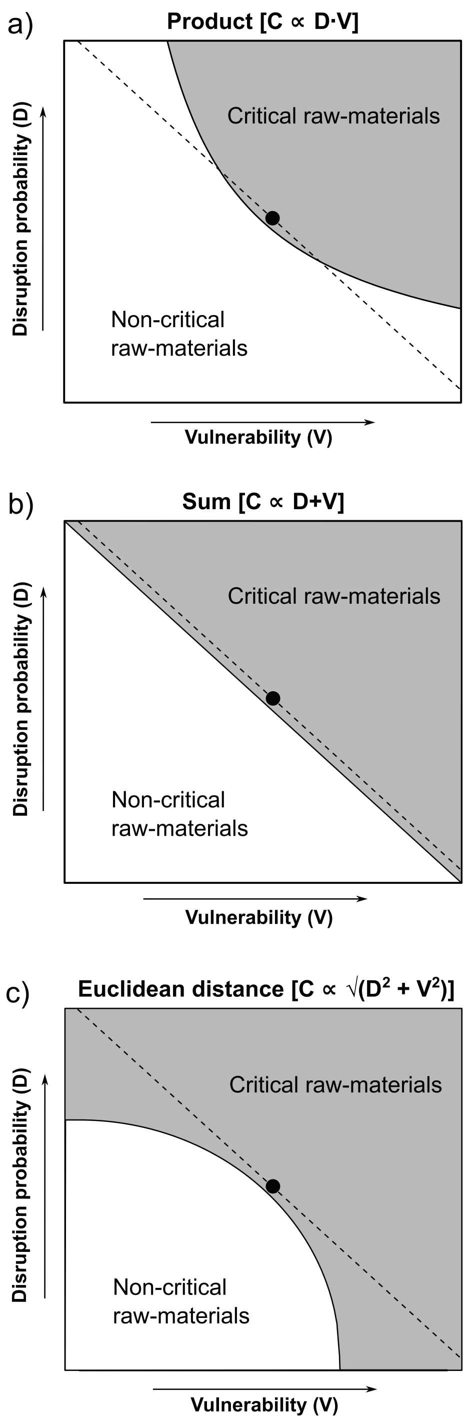

Figure 2. Schematic illustration of the effects of different calculation methods for overall criticality scores on the identification of critical raw materials: (a) product aggregation (correct method), (b) sum aggregation, and (c) Euclidean distance method. The boundaries between the fields of critical and non-critical raw materials represent iso-criticality contours, and are designed to go through the centre of the diagram. Note how the field of critical raw materials is extended towards the axes by the sum (b) and Euclidean distance (c) methods. The broken line represents the hypothetical case of perfect negative correlation between D and V, and the central dot the material with the highest true criticality score on this line. Its relationship with the iso-criticality contours indicates how criticality order is affected by the different aggregation methods, as discussed in more detail in the main text. Figure inspired by Glöser *et al* (2015).

Last but not least, it is worth noting that the incongruity between the shape of the criticality region in standard criticality plots with the shape of contours of constant criticality may also cause undesirable effects (see Glöser *et al* 2015). This is illustrated in figure 3 for the data published by the EU Commission (2014). It shows both the 'critical' region, as well as contours of constant criticality (product of the two axes). The main result is that materials with the same criticality as the lower left-hand corner of the square 'critical' region are not included in the list of critical materials. These are: baryte, Mo, Sn, natural rubber, V and bauxite. This is particularly interesting for V and natural rubber, both of which clearly have higher criticality values than borate, which is included in the list. Similarly, materials with very high criticality values (greater than those of Sb, magnesite and W), but plotting close to the upper left-hand corner of the criticality field, would not be included in a list of critical raw materials.

To briefly summarise this sub-section, not a single study complies even with the simplest formal requirements for the identification and ordering of critical raw materials set by classical risk theory. This is in accordance with the results of Glöser *et al* (2015).

3.3. Addition or omission of assessment dimensions

Another important aspect to consider when evaluating any criticality assessment is whether it honours the 2D structure dictated by (2) to (5). That is, whether it assesses both event probability and vulnerability (consequences) for the raw materials under consideration, and aggregates them in a suitable manner.

While most studies comply with the basic requirement of two-dimensionality (see section 3.1, Glöser *et al* 2015), some do not. Interestingly, deviations occur in both directions: either, additional dimensions are introduced, or one of the two original dimensions, usually vulnerability, is omitted. An excellent example for the introduction of additional dimensions is provided by the methodology of Graedel *et al* (2012), who introduce an additional 'environmental implications' dimension. Examples using only a single dimension are provided by the British Geological Survey's 'risk list' (e.g. BGS 2015), the JRC report (Moss *et al* 2013) and the Oakdene and Hollins study (Morley and Eatherly 2013).

Two important questions need to be asked with respect to these studies: first, can the introduction or omission of dimensions be justified, and second, what is its effect on the overall results of an assessment? With respect to these questions, a clear distinction must be made between the addition and omission of dimensions. In the following, we briefly discuss both cases in the context of specific examples.

As an example for the addition of a dimension, we choose the methodology of Graedel *et al* (2012). In the authors' own words, the additional 'environmental implications' axis is meant to 'depict the environmental burden of the various metals', i.e. the *ongoing* environmental effects due to their extraction and use. There are some major conceptual issues with the introduction of this aspect as an additional equivalent axis in criticality assessments.

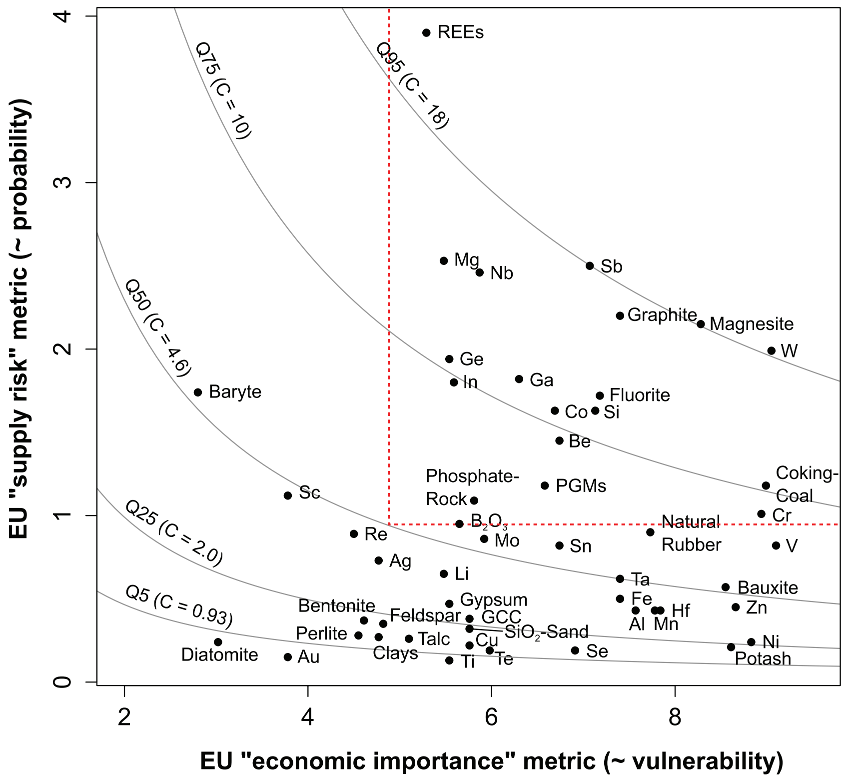

Figure 3. Re-plot of the criticality plot used in the EU study (EU Commission 2014) to identify critical raw materials, showing both the boundary of the criticality field used by the EU (broken line), and contours of constant criticality calculated as the product of the two assessment dimensions (i.e. D * V; grey lines). Contours show the 5th, 25th, 50th, 75th, and 95th quantiles of the distribution of overall criticality scores, as labelled. Plot data from EU Commission (2014), inspired by Glöser *et al* (2015).

For instance, the environmental damage caused by the ongoing extraction of a raw material forms part of its regular extraction costs, in the broadest sense of the term. As a normally externalised cost that is not represented in the price of a commodity, it can in principle be seen as equivalent to a risk (which is also an expectation value of a cost, see section 3.1), although it might be expressed in different units. Thus it would be on the same level as the (economic) criticality assessed via the other *two* axes of Graedel *et al*'s (2012) assessment, and should not be mixed with these axes, but only with the final results for economic criticality, provided a suitable harmonisation between economic criticality and environmental risk can be achieved.

Unfortunately, the authors do not provide a justification for their decision to consider environmental implications as an equivalent third dimension. In combination with the use of the Euclidean distance method for the calculation of the overall criticality score (see above) this causes a further distortion of their results, and even greater limitations to the usefulness of their methodology.

The omission of one of the original dimensions can also have a number of unfortunate effects on the overall results. To illustrate this, consider the case where the vulnerability dimension is omitted (e.g. BGS 2015). In this case, the overall result depends greatly on the nature of the correlation between disruption probability and vulnerability. If it is strongly positive, then materials with a high disruption probability will also have a high vulnerability, and therefore a high criticality. Identification and ordering of critical materials would thus be (mostly) correct, even if only one dimension were used.

If there is no correlation, however, then a significant proportion of all materials with a high disruption probability will not be critical because they will have low vulnerability. In such a case, the expected error rate of an assessment would be on the order of tens of percent, and the result little better than a random assignment of criticality scores.

Further, if the correlation is negative, then the materials with the highest disruption probability will not include those with the highest true criticality, since these materials would lie around the middle of the range in disruption probabilities (see figure 2 and corresponding discussion in section 3.2 above). Such a result might again be considered 'worse-than-useless' (see section 3.2).

An important general note from this as well as the previous sub-section must be that the outcome for any methodology not following the prescriptions from risk theory greatly depends on the covariance structure of the data. This issue is greatly exaggerated when assessments fail to include one of the two dimensions, vulnerability or disruption probability. As discussed in detail by Buijs and Sievers (2011a, 2011b), the case of negative correlation, which usually results in the worst outcomes of assessments using non-ideal methodology, might actually be expected for the global raw materials market.

3.4. Indicator choice and aggregation

In order to ensure the logical coherence and applicability of criticality assessments, the indicators used for the vulnerability and disruption probability scores, as well as their aggregation should fulfil the following general conditions in compliance with the accepted methodology for quantitative risk assessments (see Cox 2009):

(1) They should have empirically demonstrable, statistically significant relationships with the quantity of interest; and
(2) Be aggregated and weighted in a way that reflects these relationships. Corrections should be made where indicators are correlated, to avoid positive bias.

To fulfil both conditions, assessments would need to include an evaluation of indicator quality and scaling using empirical data. Additionally, an evaluation of causal and probabilistic links between different events should be performed.

Unfortunately, none of the recent assessments includes an in-depth analysis of suitable empirical data to verify indicator choices, or inform the weightings/scalings used in their aggregation. Instead, the particular choices and weightings generally reflect the subjective opinions of the authors (see Erdmann and Graedel 2011, Graedel and Reck 2015).

The absence of empirically grounded assessments is by no means evidence for the absence of suitable input data. For instance, a study by Gleich *et al* (2013) showed that the historic evolution of market prices of mineral raw materials is reasonably well explained by combinations of different explanatory variables (including HHI) whose values are freely available from various sources for at least the past 30 years. Crucially, the exact combinations of explanatory variables, and their respective weightings differ between elements, likely as a function of different market structures. Based on these results, Gleich *et al* (2013) suggested that the same is to be expected for certain aspects of criticality assessments, particularly the disruption probability (supply risk) dimension, since prices are to some degree a reflection of the scarcity of a raw material (see Gleich *et al* 2013). This means that the generalised approach used in criticality assessments, where the same indicator weightings are used for all elements, is probably a gross oversimplification that has little, or nothing, to do with reality.

3.5. Summary of evaluation

The detailed evaluation presented in the previous sub-sections illustrates that virtually all available studies on raw material criticality suffer from at least one, if not several, serious methodological short-comings that highly limit their applicability. In particular:

(1) All studies ignore the fact that there is a spectrum in the severity of the events of interest (effect size), with variable consequences, and accordingly focus on a single type of event only, without considering the effects of this simplification.
(2) Even the simplified studies do not comply with the formal requirements set by risk theory for the ordering and identification of critical raw materials. Depending on the exact methods used, as well as the nature of the correlation between disruption probability and vulnerability, 'worse-than-useless' outcomes sensu Cox (2009) may result.
(3) Some studies introduce additional dimensions, or omit the vulnerability dimension from the assessment. This can again bring about 'worse-than-useless' results, depending on data structure.
(4) Neither the choice of indicators, nor their aggregation and scaling are informed by empirical data on actual market behaviour, but rather by the subjective opinions of individual authors. Available empirical data suggests that these arbitrary procedures are probably not adequate.

Profound conceptual issues are apparent from these shortcomings. They suggest that many authors are aware of neither the fundamental equivalence of criticality and risk (see Glöser *et al* 2015), nor the true meaning of risk itself. This impression is corroborated by recent reviews (Graedel *et al* 2015, Dewulf *et al* 2016) whose authors are clearly aware of the work of Glöser *et al* (2015), but neglect to discuss its fundamental implications. Many of the shortcomings of recent criticality assessments outlined above could have been avoided with a sound understanding of risk theory.

Even without reference to risk theory, both the lack of conceptual unity between current studies, as well as their universal failure to take into account relevant empirical data, provide clear indications for the deep philosophical issues present in the field.

Despite the evident shortcomings of available work on raw material criticality, we note that criticality assessments are nevertheless important. This is for two major reasons. First, even inaccurate assessments raise awareness about the general issue of raw material supply security, a topic that until recently had mostly been ignored in western countries. Second, assessments can be amended to better reflect reality using the principles set out above. In particular, they should be based on empirical evidence, include logically coherent risk models, and be compatible with risk theory (see Cox 2009).

However, the necessary increase in complexity might pose significant challenges for assessors. Specifically, it will require extensive collaboration between different disciplines along the value chain. Geoscientists, mining and mineral process engineers, materials scientists, metallurgists, and economists all have a rôle to play. While current assessments should be regarded with healthy scepticism, improved assessments resulting from such future work would certainly be useful tools for political and industrial decision makers.

4. Other important aspects of risk in global metal markets

Because criticality assessments focus exclusively on single raw materials, they necessarily ignore some of the interconnected nature of global markets and the many complexities,

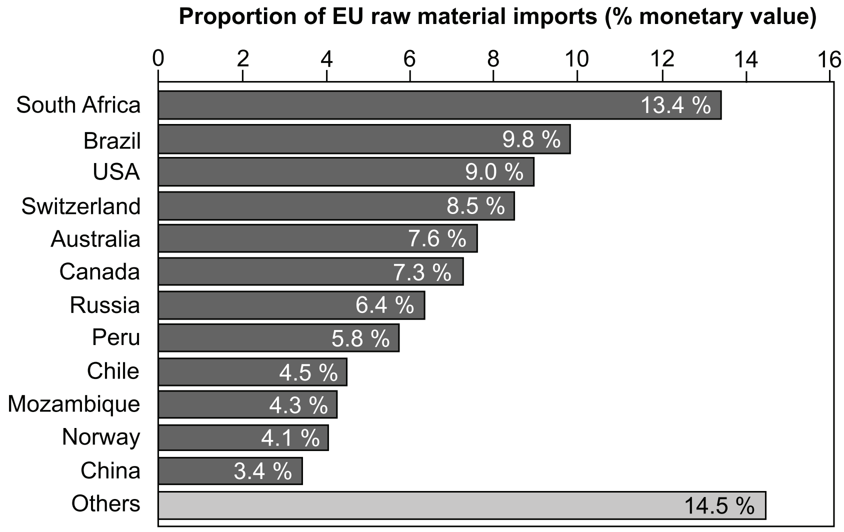

Figure 4. Contribution of different countries to overall raw material imports to the EU in 2011/2012, as a percentage of total monetary value of non-fuel raw material imports. Import data taken from EU Commission (2014), and price data for individual raw materials taken from Kelly and Matos (2014). Note that the prominent position of Switzerland is due to its role as a major trading hub for gold and other precious metals. It does not itself produce significant amounts of primary raw materials. The other countries on the list, however, are all major primary producers for various commodities.

and risk factors, arising from it. This section is devoted to a brief introduction and discussion of some of these risk factors, and their potential effects on national or multi-national economies. In particular, we focus on the following aspects:

(1) Concentration of the production/supply of not one, but *several or many* different raw materials in a small number of countries;
(2) Interdependence of the production cycles of various industrial and minor metals;
(3) Growing energy requirements due to decreasing ore deposit quality.

Each will be treated in a separate sub-section. We note that some or all of these aspects are often discussed in criticality assessments, too (e.g. EU Commission 2014, Gutzmer and Klossek 2014, Angerer *et al* 2016), but due to their nature are inherently difficult to incorporate into the material-by-material approach of these assessments.

4.1. Concentration of the supply of many raw materials in few countries

Many studies highlight the current status of China as the largest producer of many metals as a major risk factor in global markets (e.g. EU Commission 2014). The general fear is that a dependence on a small number of countries for many raw material imports increases the probability of wide-ranging supply disruptions in the same way it supposedly does so for single raw materials (see section 2.3). But is this fear justified?

As a first point, we note that the fact that China is the main producer of many mineral raw materials does not necessarily imply its dominance for the import balance or consumption of other countries. This is exemplified by the country-specific total values of raw material imports to the EU (figure 4). Clearly, China is only a minor supplier of raw materials to the EU. While no single other country has a strongly dominating share in EU imports, South Africa, the USA, Brazil, (Switzerland), Australia and Canada are clearly the EU's most important trading partners for raw materials. This is not an unexpected result given the strong historic ties between Europe and these countries, as well as their well-developed mining industries based on a world-class natural resource base.

The probable reason for China's relatively unimportant role as a supplier to the EU (and other countries) is its dual role as both the largest producer and the largest *consumer* of raw materials worldwide (Streifel 2006, Roache 2012). Only for some of the minor raw materials (e.g. rare earth elements, antimony) does China play an important role as an exporter. However, due to their comparatively low total value and specialised applications, these materials only have a very small effect on overall import balances.

Today, most of China's primary production is consumed internally, and only exported as part of intermediate or finished industrial products. In terms of *total* imports to the EU, China is indeed its largest trading partner (it is second in terms of exports from the EU, Eurostat 2016). This current state of affairs is a direct result of Chinese efforts to integrate manufacturing capacities along the value chain, and concentrate them in China, starting with the production of raw materials, and then moving on to industrial products. Given this trend, perhaps greater attention should now be paid to the EU's reliance on China for the provision of certain industrial products (e.g. consumer electronics) rather than specific raw materials.

The emergence of China as the largest consumer of mineral raw materials has had other important consequences for global markets. In particular, the greater competition for resources has resulted in increased overall demand, and consequently significant increases in raw material prices (Streifel 2006, Roache 2012). This general effect is quite different from the

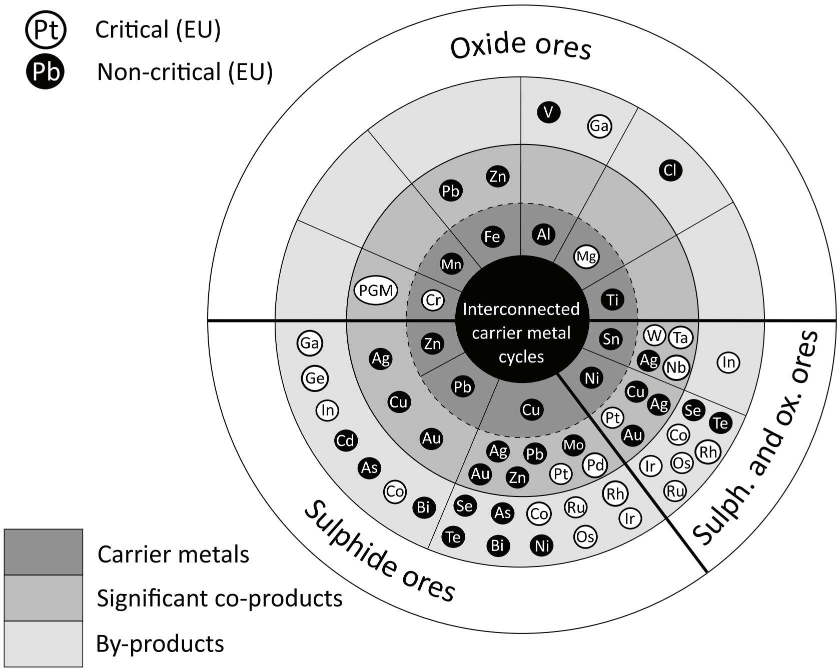

Figure 5. The metal wheel—a schematic illustration of the geological relationships between different industrial metals and their co- and by-products. Modified from Verhoef *et al* (2004). Criticality assignments according to the study of the EU Commission (2014).

feared supply restrictions due to individual producing countries discussed in many criticality assessments. Despite its evident importance, we are not aware of a systematic assessment of the risk posed to global raw material markets, and western economies in particular, by the rapid growth of other emerging economies, e.g. India.

A related point is the loss of primary production infrastructure and the related know-how to other countries, e.g. due to competitive disadvantages (price-dumping etc.). Since this makes the (re-)establishment of significant local production much harder, it has the potential to greatly exacerbate supply disruptions, should they occur. This poses a risk that must not be underestimated (Reuter 2016).

4.2. Interdependence of metal-cycles

While the separate treatment of different raw materials in criticality assessments suggests that each material is produced independently of all others, this is not generally the case. Many metals tend to occur together in the same geological environments due to similarities in their geochemical behaviour. Good examples of such groups are Pb–Zn–Ag, U–Th-Rare Earths and Cr–Ni–Platinum group elements (see Robb 2005). The metal wheel of Verhoef *et al* (2004) provides an overview of the relationships between a number of different major and minor metals (figure 5).

While many mines produce one metal as the main product, they often also produce one or several co- or by-products. The difference between these two categories is that co-products, due to their comparatively high value, still enact an influence on mining decisions, while by-products, due to their low value, do not (Campbell 1985). To an extent, the co- and by-products may help to maintain the profitability of a mining operation.

The main issue, particularly with the co-production of different metals, is that movements in the market of one metal can have significant effects on that of another, and vice versa. Consider the simple example where for some reason the use of Pb in all branches of industry in the EU is banned, resulting in a significant drop in Pb-prices. Because virtually all Pb mines are also Zn mines, and vice versa, each element being an important co-product of the other (Reuter *et al* 2015; see figure 5), this would lead to a drop in Zn production due to mine closures. Therefore, the end result would be an increase in Zn prices, an outcome that would have been difficult to predict without knowledge about the close geological link between these two metals. Similar linkages exist between other metals (figure 5). The nature of these links is not incorporated into criticality assessments in a meaningful way. The percentage of a specific element produced as a 'companion' used in some assessments (Graedel *et al* 2015, Nassar *et al* 2015) does not provide useful information on the relevant peculiarities of its specific market, or the markets it is related to (e.g. the Pb–Zn linkage).

For by-products, the relationship is more one-sided. In particular, the production rate of the corresponding main products sets a relatively strong limit to their supply (e.g. Campbell 1985). This might result in supply restrictions once this limit is reached (e.g. Frenzel *et al* 2015, Nassar *et al* 2015, Lovik *et al* 2016). However, this is an element-specific problem, since the by-products do not affect the corresponding main products. As such, it is easily incorporated into criticality assessments. Only the actual determination of true supply potentials might present some difficulties due to data availability issues (e.g. Frenzel *et al* 2015, 2016). Supply potentials, however, are not considered in any criticality assessment. As mentioned earlier, other issues are often correlated with the by-product status of some raw materials, e.g. intransparent trading and price

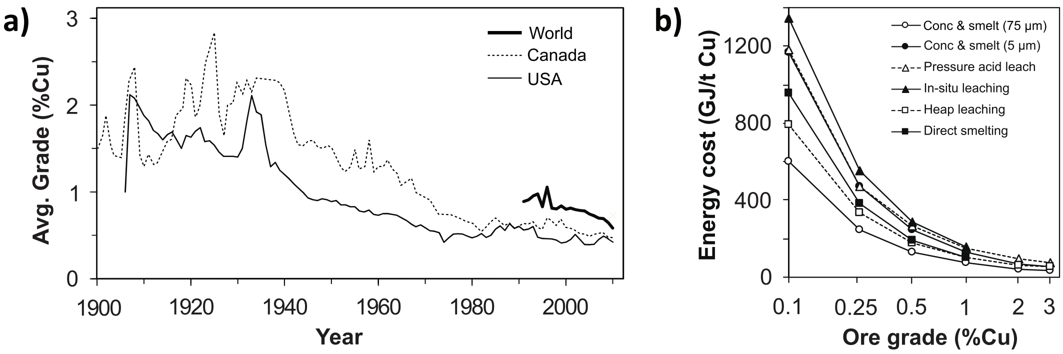

Figure 6. Decreasing ore quality and its effects on unit energy costs of production: (a) decline in mean copper ore grades over the last century for Canada, the USA and the World (data from Mudd *et al* 2013), and (b) increase in energy costs with decreasing ore grade (modified from Norgate and Jahanshahi 2010).

formation, and high concentration of supplies (Wellmer *et al* 1990, Frenzel *et al* 2015). These issues are mostly a function of the small market size for these materials (see Frenzel *et al* 2015). Small market size (in terms of monetary value of total consumption) can be seen as a first order indicator for vulnerability. This means that for most of these materials, overall criticality is probably small, even if disruption probability is relatively high.

4.3. Decreasing ore quality

It is well established that access to mineral resources is becoming harder as time progresses (Mudd 2009, Mudd 2010, Northey *et al* 2014). Following the 'best first' principle (Hall and Kiltgaard 2012), the best ore deposits (near-surface, high-grade, close to major infrastructure) are already found and exploited. This has two consequences: First, new ore deposits become increasingly hard to find, and second, if found, the resources they contain are generally harder to access. The latter might be for several reasons, e.g. greater depth below surface, lower grade, or remoteness of location. A good example of this general trend is provided by the steady decline in mean copper ore grades over the last century (e.g. Mudd *et al* 2013, figure 6). As eloquently demonstrated by Mudd (2010) this is both a geological as well as an economic inevitability.

Both effects, but particularly the second, i.e. harder access, tend to increase energy costs per unit mass of material (e.g. Calvo *et al* 2016), although this may be counteracted to some degree by changes in extraction technology (e.g. Swart and DeWulf 2013). This trend is particularly problematic in view of the increased non-fuel raw material needs for decentralised energy production using renewable sources (Kleijn *et al* 2011, Vidal *et al* 2013). In particular, significantly greater needs for non-fuel raw materials per unit energy produced (e.g. figure 7) could necessitate the mining of non-fuel raw materials at ever greater energy costs which would in turn increase the total demand for energy, and therefore non-fuel raw materials, creating a positive feedback loop with potentially dramatic consequences for the global economy. Note that renewables (particularly wind and solar PV) do not just have significantly higher needs in terms the industrial metals (Fe, Cu, Al) and concrete (figure 7), but also many of the rarer elements (e.g. Ag, Ge and Sn, see Kleijn *et al* 2011). However, in terms of total energy requirements, the industrial metals are much more significant.

Already, the post-mine processing and refining of non-fuel raw materials to metal and mineral products account for about 12% of global energy consumption (EIA 2016). The energy consumption of mining is not included in this figure, but is probably of a similar order of magnitude (e.g. mining accounted for c. 11% of total energy demand in South Africa in 2012, DoEZA 2013).

This systemic problem cannot be studied through the lens of criticality studies, but rather requires an integrated approach to reach an assessment of the potential severity of the issue. While a number of authors have previously pointed out this problem (Kleijn *et al* 2011, Vidal *et al* 2013, Fizaine and Court 2015), a detailed study of its potential magnitude is still lacking.

4.4. Summary

In summary, there are a number of complex risk factors present in global markets, which are not, and sometimes cannot, be captured in criticality assessments as long as these are done on a material-by-material basis. The most pressing of these issues is probably the close link between the renewable energy sector and increased future raw materials demand, with potentially disastrous consequences for the global economy. As emphasised by Liebig as early as 1866, 'civilisation is the economy of power' (cited in Jevons 1866). Without a source of abundant cheap energy (Liebig went on to call coal England's source of power), human civilisation would not have progressed as rapidly as it did, nor will it be able to continue this progress into the future (Jevons 1866). If an increasing proportion of total energy production needs to be devoted to the extraction of primary raw materials, this main enabler of technological progress is under serious threat.

5. Alleviation of risks

While the previous sections focused primarily on the nature of the different risks present for consumers in the global raw materials markets, this section is dedicated to measures proposed for the mediation of these risks. Given the limited applicability of current criticality studies, however, it is not actually clear which

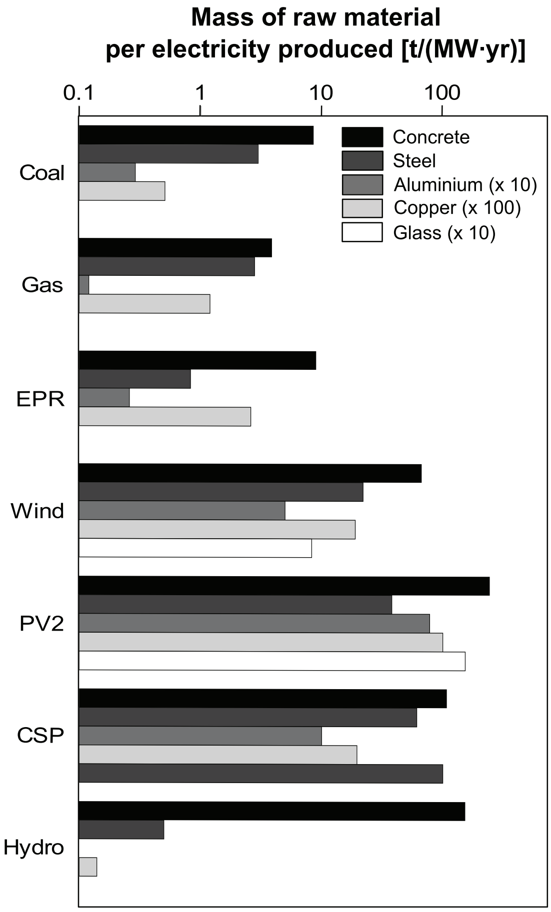

Figure 7. Raw material requirements per unit energy produced for different sources of electricity. EPR—evolutionary pressurised reactor; PV—photovoltaics, CSP—concentrated solar power. Note the significantly higher demand for industrial metals and concrete for electricity produced from renewable sources. Also note that coal and gas do not include contributions from fuel extraction. In terms of total greenhouse gas emissions, renewables still do significantly better than fossil sources of electricity (Kleijn *et al* 2011). Data recalculated from sources in Vidal *et al* (2013).

individual risks are the greatest. Therefore, this section must necessarily be of a preliminary character. It will focus on those measures discussed most frequently in the recent debate.

Mediation measures can be subdivided into two main groups according to which part of the overall risk they focus on: they can reduce either disruption probability, or vulnerability. In the following sub-sections, different proposed measures are introduced and their main effects on overall risk are discussed.

5.1. Reduction of disruption probability

One traditional strategy to lower disruption probability due to high import dependence is diversification of supplies. Companies usually diversify their supplies when the main supplier either increases prices or struggles with delivery problems, or proves to be no longer trustworthy. Diversification can be a good strategy to mitigate the influence of a quasi-monopoly on a market, and to lower the probability for supply disruptions. However, sometimes diversification is not easy to achieve due to unfavourable market conditions such as limited opportunities to diversify (only a small number of suppliers exist) or when investments to develop alternative resource projects demand a long time scale as is the case in the REE market (de Boer and Lammertsma 2013, Klossek *et al* 2016). Other factors that can also have a negative effect on diversification efforts are general political instability of producing countries, civil wars and underdeveloped or unsecure trade routes.

Other than the political measures focusing on international relations and diversification of imports, another possibility for the reduction of disruption probability is to increase the share of domestic production in the total consumption of the most critical raw materials. Policymakers can directly influence the commercial success of mining operations through the amount of taxation, environmental laws or export duties. Depending on the country, political decisions can become critical factors for the mining business (Arsel *et al* 2014). Encouraging greater recycling rates within the framework of the so-called circular economy (Yuan *et al* 2006, Anderson 2007, EU Commission 2015), again to increase the share of domestic production in total consumption, might also be a viable option to decrease the probability for supply disruptions from international political conflicts.

However, once primary production infrastructure and the related know-how required for its operation has been lost from a country or region, the (re-)establishment of primary production is much more difficult (Reuter 2016). Therefore, policy should also focus on avoiding such losses, as they would inadvertently stifle—or at least delay – any efforts to stimulate domestic supply chains.

Concerns about undisturbed access to land-based raw materials also opened up discussion for the potential future exploitation of marine resources (e.g. EU Commission 2012, Greenpeace 2013, OECD 2016). At the moment, deep-sea mining is technologically challenging and commercially not viable. However, if raw material prices should increase again in the future, deep sea deposits of polymetallic nodules and manganese crusts rich in cobalt, copper, nickel, manganese and even platinum, tungsten and rare earths could become valuable strategic resources (Hein *et al* 2013). This is why several nations as well as exploration companies have secured exploration licenses from the International Seabed Authority (International Seabed Authority 2014). These exploration activities could lead to the exploitation of deep-sea resources in the next two decades. However, industrial-scale mining on the seafloor remains a controversial issue. Critics emphasise the limited understanding of the potential impact on deep-sea ecosystems, requiring further scientific study (Greenpeace 2013). Besides these ecological concerns, the higher energy requirements for the exploitation of seabed resources again feeds into the decreasing quality of, and harder access to, natural resources discussed earlier.

5.2. Reduction of vulnerability

Besides measures to reduce disruption probability, a number of different options also exist to reduce the vulnerability to supply shortfalls. For instance, both the effects of transient price-hikes and supply shortfalls can be mediated by keeping stockpiles of the required raw materials covering a few months, or even a year, of consumption. This can be done at either the company or national level. Japan and South Korea, for instance, maintain national stockpiles of raw materials they consider critical for their manufacturing industries (JOGMEC 2014, KORES 2013).

Another option to reduce vulnerability, particularly the economic impact of price-hikes, is to reduce the amount of critical materials used in the final products. This can either be done through more efficient product design or production processes, or through increased substitution where viable substitutes exist for critical raw materials (e.g. as for Co in permanent magnets in the 1970s, see Buijs *et al* 2011a, 2011b).

Of course, policymakers can have a direct influence on these processes by encouraging research into resource efficiency and material substitution. However, individual companies will also have to invest into their own research and development programmes, particularly if they are more seriously affected by the risks related to certain raw materials than the corresponding national economies. As mentioned in the beginning, criticality is dependent on the specific consumer. Raw materials critical for one or two companies might not be critical for a whole country.

5.3. The long term

It is also important to re-iterate that one of the main raw material-related long-term risks to the global economy is the prospect of increasing unit energy cost of production at increased demand. There are many options to address this risk. For instance, the effect might be counteracted to some degree by the introduction of new mining and beneficiation processes that are considerably more energy efficient than conventional options. There is a significant potential for such reductions in the mining industry (e.g. with more efficient crushing and grinding equipment; Napier-Munn 2015).

An alternative strategy is to deploy (new) means of less resource-intensive, zero-carbon energy production beyond the current drive for renewables. There are a few options, e.g. fossil fuels with carbon capture and storage, nuclear fission, and perhaps nuclear fusion (Brook 2012, Cowley 2016). While fission power is already available for large-scale deployment, the other two options still require considerable future development (Boot-Handford *et al* 2014, Lopez Cardoso *et al* 2016).

Movement towards a 'circular economy' (Yuan *et al* 2006, Anderson 2007, EU Commission 2015) where recycling rates for all non-consumable raw materials approach 100% is another important option. It should be noted, however, that perfect recycling is a physical impossibility (e.g. Reuter *et al* 2006, Amini *et al* 2007, Reck and Graedel 2012), and, therefore, that the mining of primary resources will probably never cease completely. A particular problem is that the recycling of ever more complex products will also require increasing amounts of energy per unit mass of raw material, just as for primary ores (Reuter *et al* 2006).

Nevertheless, increased rates of recycling may serve to take some of the strain from the global primary production system. This could be complemented by reductions in raw material consumption, e.g. through better raw materials efficiency in industrial products, or substitution of the most energy-intensive raw materials by less energy-intensive ones. Materials research has an important role to play in these strategies.

The best results would evidently be achieved by a favourable combination of all of the above strategies. However, while there are obvious economic incentives for their adoption, targeted government policies will be needed to accelerate the process. Particularly in areas where basic research is required to lay the scientific foundations for new and more efficient processes, as well as to demonstrate their viability, government funding is often an essential ingredient. The importance of the adoption of suitable policies can thus hardly be understated.

5.4. The role of applied physics

Since this article is chiefly intended for an applied physics audience, we would now like to highlight some of the many areas in which physicists can make an important contribution to the mediation of risks related particularly to escalating unit energy costs of production for many raw materials. Specifically, the following areas of physics are of interest:

(1) **Solid state physics**: first, for research into the substitution of the most energy-intensive and critical primary resources in industrial materials; second, to develop new materials with the right properties required for alternative modes of power production (e.g. in solar cells or nuclear reactors); and third, for the development of more efficient grinding techniques, which will require a detailed understanding of the physical controls on rock breakage, and innovative thinking regarding energy-efficient fracturing (Napier-Munn 2015).
(2) **Nuclear physics**: for the development of alternative modes of power generation (fusion/fission), to supplement renewable sources.
(3) **Surface physics/fluid dynamics**: for the modelling and improvement of current mineral separation technologies, particularly froth flotation.
(4) **Exploration geophysics**: for the development of new geophysical/remote-sensing techniques to improve the success rate, and therefore the efficiency, of mineral exploration efforts.

Of course, this list is by no means exhaustive. Nevertheless, it still illustrates the relevance of many areas of applied physics to the complex problem at hand.

6. Summary and conclusions

After a detailed evaluation of the methodology of current criticality assessments, as well as the more complex resource-related challenges facing the global economy, we come to the following conclusions:

First, current assessments of raw material criticality are flawed in several fundamental ways. This is mostly due to a lack of adherence to risk theory, and highly limits their applicability. Many of the raw materials generally identified as critical, particularly those produced as by-products and only employed in niche, albeit high-tech, applications, are probably not critical.

Second, the flaws of current assessments do not mean that the general issue of supply security of raw materials can be ignored, but rather, that new assessments are urgently needed. These assessments should follow the principles outlined in section 3 of this review. In particular, they should be based on empirical evidence, include logically coherent risk models, and be compatible with risk theory (see Cox 2009).

Third, the results of new and improved assessments will likely include more of the traditional industrial metals in the list of critical materials, particularly those related to steelmaking (Fe ore, coking coal, Cr, Ni), and power infrastructure (Cu). Increased and sustained public investment in the raw materials sector is needed to ensure the future supply security of important raw materials. This is particularly true for those societies that are large consumers but have shown rapidly decreasing public and private investment in the resource sector (Western Europe, the United States).

Fourth, the greatest challenge in the resource sector for the longer term is to counteract the escalation of unit energy costs of production, which is inevitable if conventional production processes continue to be applied to resources of ever harder accessibility—from both primary and secondary sources. This issue is particularly pressing due to its close link to the renewable energy revolution which is set to dramatically increase the demand for raw materials per unit energy produced (Vidal *et al* 2013). Due to its multi-faceted nature, the close link between energy and resource efficiency is not captured in any of the current criticality assessments.

Particularly the solution to this last challenge will require coordinated policy action, as well as the collaboration of scientists from many different fields—with physics, as well as the materials and earth sciences in the lead.

Acknowledgments

We gratefully acknowledge the invitation of Mr Dean Williams to contribute this review to the special issue on critical raw materials. We are also indebted to three anonymous reviewers for their insightful comments and suggestions, which helped to greatly improve the original manuscript. The continued support of our institutions, without which the completion of this work would not have been possible, is very much appreciated.

References

Achzet B and Helbig C 2013 How to evaluate raw material supply risks—an overview *Resour. Policy* **38** 435–47

Ackerman F and Heinzerling L 2002 Pricing the priceless: cost-benefit analysis of environmental protection *Univ. Penn. Law Rev.* **150** 1553–84

Amini S H, Remmerswaal J A M, Castro M B and Reuter M A 2007 Quantifying the quality loss and resource efficiency of recycling by means of exergy analysis *J. Clean. Prod.* **15** 907–13

Anderson M S 2007 An introductory note on the environmental economics of the circular economy *Sustain. Sci.* **2** 133–40

Angerer G, Buchholz P, Gutzmer J, Hagelüken C, Herzig P, Littke R, Thauer R K and Wellmer F-W 2016 *Rohstoffe für die Energieversorgung der Zukunft. Geologie–Märkte–Umwelteinflüsse* (Munich: Acatech)

Arsel M, Mena C, Pellegrini L and Radhuber I 2014 *Property Rights, Nationalisation and Extractive Industries in Bolivia and Ecuador Conflicts Over Natural Resources in the Global South—Conceptual Approaches* ed M Bavinck *et al* (London: CRC Press) pp 109–28

Benjamin J R and Cornell C A 1970 *Probability, Statistics, and Decision for Civil Engineers* (New York: McGraw-Hill)

BGS 2015 *Risk List* (Nottingham: BGS)

Boot-Handford M E *et al* 2014 Carbon capture and storage update *Energy Environ. Sci.* **7** 130–89

Bradshaw A M, Reuter B and Hamacher T 2013 The potential scarcity of elements for the Energiewende *Green* **3** 93–111

Brook B W 2012 Could nuclear fission energy, etc., solve the greenhouse problem? The affirmative case *Energ. Policy* **42** 4–8

Buijs B and Sievers H 2011a Critical thinking about critical materials. Assessing risks related to resource security *CIEP-BGR Briefing Paper* (www.polinares.eu)

Buijs B and Sievers H 2011b Resource security risks in perspective—complexity and nuance *POLINARES Working Paper No* 33 (www.polinares.eu)

Buijs B, Tercero Espinoza L A and Sievers H 2012 Limits to the critical raw materials approach *Proc. Inst. Civil Eng. Waste Res. Manag.* **165** 201–8

Calvo G, Mudd G, Valero A and Valero A 2016 Decreasing ore grades in global metallic mining: a theoretical issue or a global reality *Resources* **2016** 36

Campbell G A 1985 The role of co-products in stabilizing the metal mining industry *Resour. Policy* **11** 267–74

Cordier D J 2011 Rare earths *USGS Mineral Commodity Summaries* (Reston, VA: USGS) pp 128–9

Cowley S C 2016 The quest for fusion power *Nat. Phys.* **12** 384–6

Cox L A 2008 What's wrong with risk matrices? *Risk Anal.* **28** 497–512

Cox L A 2009 *Risk Analysis of Complex and Uncertain Systems* (New York: Springer)

de Boer M A and Lammertsma K 2013 Scarcity of rare earth elements *ChemSusChem* **6** 2045–55

Dewulf J, Blengini G A, Pennington D, Nuss P and Nassar N T 2016 Criticality on the international scene: quo vadis? *Resour. Policy* **50** 169–76

DoEZA 2013 *Draft 2012 Integrated Energy Planning Report. Annexure A—Technical Report. Part 1: Demand Modelling Report* (Republic of South Africa: Department of Energy)

EIA 2016 *International Energy Outlook* (Washington, DC: U.S. Energy Information Administration)

Erdmann L and Graedel T E 2011 Criticality of non-fuel minerals: a review of major approaches and analyses *Environ. Sci. Technol.* **45** 7620–30

EU Commission 2010 *Critical Raw Materials for the EU. Report of the Ad-hoc Working Group on Defining Critical Raw Materials* (Brussels: EU Commission)

EU Commission 2012 *Blue Growth. Opportunities for Marine and Maritime Sustainable Growth. COM (2012) 494 Final* (Brussels: EU Commission)

EU Commission 2014 *Critical Raw Materials for the EU. Report of the Ad-Hoc Working Group on Defining Critical Raw Materials* (Brussels: EU Commission)

EU Commission 2015 *Closing the Loop—An EU Action Plan for the Circular Economy. COM (2015) 614 Final* (Brussels: EU Commission)
Eurostat 2016 *Extra-EU Trade by Partner* (Luxembourg: Eurostat) *Last update: 01-04-2016*
Fizaine F and Court V 2015 Renewable electricity producing technologies and metal depletion: a sensitivity analysis using the EROI *Ecol. Econ.* **110** 106–18
Frenzel M, Tolosana-Delgado R and Gutzmer J 2015 Assessing the supply potential of high-tech metals—a general method *Resour. Policy* **46** 45–58
Frenzel M, Ketris M P, Seifert T and Gutzmer J 2016 On the current and future availability of gallium *Resour. Policy* **47** 38–50
Gleich B, Achzet B, Mayer H and Rathgeber A 2013 An empirical approach to determine specific weights of driving factors for the price of commodities—a contribution to the measurement of the economic scarcity of minerals and metals *Resour. Policy* **38** 350–62
Glöser S, Tercero Espinoza L, Grandenberger C and Faulstich M 2015 Raw material criticality in the context of classical risk assessment *Resour. Policy* **44** 35–46
Graedel T E, Harper E M, Nassar N T, Nuss P and Reck B K 2015 Criticality of metals and metalloids *Proc. Natl Acad. Sci.* **112** 4257–62
Graedel T E and Nuss P 2014 Employing considerations of criticality in product design *JOM* **66** 2360–6
Graedel T E and Reck B K 2015 Six years of criticality assessments. What have we learned so far? *J. Ind. Ecol.* **20** 692–9
Graedel T E *et al* 2012 Methodology of metal criticality determination *Environ. Sci. Technol.* **46** 1063–70
Greenpeace 2013 *Deep Seabed Mining. An Urgent Wake-Up Call to Protect Our Oceans* (Amsterdam: Greenpeace)
Grubb M and Newbery D 2007 Pricing carbon for electricity generation: national and international dimensions (doi:10.17863/CAM.5121)
Gutzmer J and Klossek A 2014 Die Versorgung mit wirtschaftskritischen Rohstoffen—Eine Ursachensuche und -analyse *Strategische Rohstoffe—Risikovorsorge* ed P Kausch *et al* (Berlin: Springer) pp 61–73
Hall C A S and Kiltgaard K A 2012 *Energy and the Wealth of Nations—Understanding the Biophysical Economy* (New York: Springer)
Hedrick J B 2010 Rare earths *USGS Mineral Commodity Summaries* (Reston, VA: USGS) pp 128–9
Hein J R, Mizell K, Koschinsky A and Conrad T A 2013 Deep-ocean mineral deposits as a source of critical metals for high- and green-technology applications: comparison with land-based resources *Ore Geol. Rev.* **51** 1–14
Helbig C, Wietschel L, Thorenz A and Tuma A 2016 How to evaluate raw material vulnerability—an overview *Resour. Policy* **48** 13–24
Hirschman A O 1964 The paternity of an index *Am. Econ. Rev.* **54** 761–2
Humphreys D 1995 Whatever happened to security of supply? Minerals policy in the post-Cold War world *Resour. Policy* **21** 91–7
International Seabed Authority 2014 *Contractors for Seabed Exploration* (Kingston: International Seabed Authority)
Jevons W S 1866 *The Coal Question; An Inquiry Concerning the Progress of the Nation, and the Probably Exhaustion of Our Coal-Mines.* (London: Macmillan)
JOGMEC 2014 *Annual Report 2014* (Tokyo: Japan Oil, Gas and Metal National Corporation)
Kelly T D and Matos G R 2014 Historical statistics for mineral and material commodities in the United States (2014 version) *USGS Data Series* **140**
Kleijn R, van der Voet E, Kramer G J, van Oers L and van der Giesen C 2011 Metal requirements of low-carbon power generation *Energy* **36** 5640–8
Klossek P, Kullik J and van den Boogaart K G 2016 A systemic approach to the problems of the rare earth market *Resour. Policy* **50** 131–40
KORES 2013 *Overview* (Seoul: Korea Resources Corporation)
Lapko Y, Trucco P and Nuur C 2016 The business perspective on materials criticality: evidence from manufacturers *Resour. Policy* **50** 93–117
Lopez Cardoso N J, Lange A G G and Kramer G J 2016 Fusion: expensive and taking forever? *J. Fusion Energ.* **35** 94–101
Lovik A N, Restrepo E and Müller D B 2016 Byproduct metal availability constrained by dynamics of carrier metal cycle: the gallium-aluminum example *Environ. Sci. Technol.* **50** 8453–61
Morley N and Eatherly D 2013 *Material Security—Ensuring Resource Availability for the UK Economy* (Chester: C-Tech Innovation)
Moss R L *et al* 2013 *Critical Metals in the Path Towards the Decarbonisation of the EU Energy Sector* (Luxembourg: Publications Office of the EU)
Mudd G M 2009 Historical trends in base metal mining: Backcasting to understand the sustainability of mining *Proc. 48th Annual Conf. of Metallurgists* (Sudbury, ON: Canadian Metallurgical Society)
Mudd G M 2010 The environmental sustainability of mining in Australia: key mega-trends and looming constraints *Resour. Policy* **35** 98–115
Mudd G M, Weng Z and Jowitt S 2013 A detailed assessment of global Cu resource trends and endowments *Econ. Geol.* **108** 1163–83
Napier-Munn T 2015 Is progress in energy-efficient comminution doomed? *Miner. Eng.* **73** 1–6
Nassar N T, Graedel T E and Harper E M 2015 By-product metals technologically essential but have problematic supply *Sci. Adv.* **1** e1400180
NRC 2007 *Minerals, Critical Minerals, and the US Economy* (Washington, DC: The National Academies Press)
NSTC 2016 *Assessment of Critical Minerals: Screening Methodology and Initial Application. Product of the Subcommittee on Critical Strategic Mineral Supply Chains of the Committee on Environment, Natural Resources, and Sustainability* (Washington, DC: National Science and Technology Council)
Norgate T and Jahanshahi S 2010 Low grade ores—smelt, leach or concentrate? *Miner. Eng.* **23** 65–73
Northey S, Mohr S, Mudd G M, Weng Z and Giurco D 2014 Modelling of future copper ore grade decline based on a detailed assessment of copper resources and mining *Resour. Conserv. Recy.* **83** 190–201
OECD 2016 *The Ocean Economy in 2030* (Paris: OECD)
OED 2016 *Oxford English Dictionary. The Definitive Record of the English Language* (Oxford: Oxford University Press)
Panousi S, Harper E M, Nuss P, Eckelman M J, Hakimian A and Graedel TE 2016 Criticality of seven specialty metals *J. Ind. Ecol.* **20** 837–53
Reck B K and Graedel T E 2012 Challenges in metal recycling *Science* **337** 690–5
Reuter M A 2016 Digitalizing the circular economy—circular economy engineering defined by the metallurgical internet of things *Metall. Mater. Trans.* B **47** 3194–220
Reuter M A, van Schaik A, Ignatenko O and de Haan G J 2006 Fundamental limits for the recycling of end-of-life vehicles *Miner. Eng.* **19** 433–49
Reuter M A, Matusewicz R and van Schaik A 2015 Lead, zinc and their minor elements: enablers of a circular economy *ERZMETALL* **68** 134–48
Roache S K 2012 China's impact on World commodity markets *IMF Working Paper* (Washington, DC: IMF)
Robb L 2005 *Introduction to Ore-Forming Processes* (Oxford: Blackwell)
Schwarz-Schampera U 2014 Indium *Critical Metals Handbook* ed G Gunn (New York: Wiley) pp 204–29

Streifel S 2006 *Impact of China and India on Global Commodity Markets—Focus on Metals & Minerals and Petroleum. Draft. Development Prospects Group.* (Washington, DC: World Bank)

Suter G W II 2006 *Ecological Risk Assessment* (New York: CRC Press)

Swart P and Dewulf J 2013 Modeling fossil energy demands of primary nonferrous metal production: the case of copper *Environ. Sci. Technol.* **47** 13917–24

Tse P-K 2011 China's rare-earth industry *USGS Open-File Report* 2011-1042

Verhoef E V, Dijkema G P J and Reuter M A 2004 Process knowledge, system dynamics, and metal ecology *J. Ind. Ecol.* **8** 23–43

Vidal O, Goffé B and Arndt N 2013 Metals for a low-carbon society *Nat. Geosci.* **6** 894–6

Wellmer F-W, Hannak W, Krauss U and Thormann A 1990 Deposits of rare metals. Raw materials for new technologies *Proc. 5th Int. Symp., Held in Hannover, Fed. Rep. of Germany, at the Federal Institute for Geosciences and Natural Resources (19–21 October 1988)* ed M Kürsten (Stuttgart: Schweizerbart) pp 71–122

Werner T T, Mudd G M and Jowitt S M 2017 The world's by-product and critical metal resources part II: a method for quantifying the resources of rarely reported metals *Ore Geol. Rev.* **80** 658–75

World Bank 2015 *Worldwide Governance Indicators* (Washington, DC: World Bank)

Yuan Z, Bi J and Moriguichi Y 2006 The circular economy—a new development strategy in China *J. Ind. Ecol.* **10** 4–8